Design of transcriptional regulatory elements for *Streptomyces*

Filipe Castro Marques
Mestrado em Bioquímica
FCUP/ICBAS
2015

Orientadora
Marta Vaz Mendes, PhD, Bolseira de Pós-doutoramento em Instituto de Biologia Molecular e Celular, Universidade do Porto

Co-orientadora
Rute Oliveira, Investigadora em Instituto de Biologia Molecular e Celular, Universidade do Porto

IBMC
INSTITUTO DE BIOLOGIA MOLECULAR E CELULAR
INSTITUTE FOR MOLECULAR AND CELL BIOLOGY

U.PORTO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR
UNIVERSIDADE DO PORTO

U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____

Há sempre qualquer coisa que está pra acontecer

Qualquer coisa que eu devia perceber

Porquê, não sei

Porquê, não sei

Porquê, não sei ***ainda.***

José Mário Branco

Agradecimentos

A realização desta dissertação foi possível apenas pela conjugação de esforços e vontades de muitos intervenientes para além do próprio autor.

Agradeço ao Professor Doutor Pedro Moradas-Ferreira por me acolher no grupo *Cellular and Applied Microbiology*.

Agradeço à Doutora Marta Vaz Mendes pela oportunidade de desenvolver o trabalho sob sua orientação, pelo seu acompanhamento ao longo do trabalho prático e pela revisão desta dissertação.

Agradeço à Rute Oliveira pelos conhecimentos metodológicos que pacientemente me transmitiu, pelo sentido crítico e o respeito pelos demais investigadores que me incutiu e pela revisão desta dissertação.

Agradeço aos restantes membros do grupo *Cellular and Applied Microbiology*, presentes e passados, por serem gentilmente prestáveis e por partilharem entusiasticamente dos mesmos interesses científicos.

Agradeço aos membros do grupo *Bioengineering and Synthetic Microbiology* por todo o apoio e incentivo na senda da biologia sintética. Agradeço aos membros do grupo *Redox Cell Signalling* pela boa convivência e companheirismo.

Agradeço, enfim, a todos aqueles que fora do laboratório me deram ânimo e mostraram interesse pelo meu trabalho.

Agradeço aos meus pais, sabendo que nunca o farei em demasia.

Abstract

Streptomyces spp. are soil-dwelling bacteria well known for the ability to produce high added value metabolites for biotechnological application. The activation of secondary metabolites silent biosynthetic gene clusters encoded in *Streptomyces* genomes has been regarded as the renaissance of natural product discovery pipeline. Synthetic biology approaches are now been adopted for metabolic (re-)design and optimization. In this context, a diverse and well-characterized molecular toolbox is essential for construction and implementation of genetic circuits. The available constitutive promoters for *Streptomyces* are scarce and do not fulfil the standards of predictability and characterization needed for synthetic biology.

In this work, publically available genome-wide transcriptomic and proteomic data of the model organism *Streptomyces coelicolor* A3(2) was used to obtain a short list of 12 stable and highly expressed genes. Further *in silico* analysis of the promoter regions of selected genes allowed to design three synthetic promoters (P1, P2 and P3). The designed promoters were characterized in *Streptomyces coelicolor* M145 and *Streptomyces lividans* 1326 in time-lapse studies, under heat-shock conditions and in defined carbon source media. This study provides proof-of-concept for a new strategy of promoter designing in *Streptomyces.*

Keywords: *Streptomyces*, RNA-seq, promoters, *in silico* analysis

Resumo

As bactérias do género *Streptomyces* são reconhecidas pela diversidade e valor dos compostos secundários que produzem. O crescente número de genomas sequenciados de *Streptomyces* tem revelado um enorme potencial de síntese de novos compostos que não são expressos em condições de laboratório. A biologia sintética é uma das metodologias possíveis para remodelar a organização do genoma de modo a produzir esses compostos. Este tipo de engenharia genética requer ferramentas moleculares bem caracterizadas. Os promotores constitutivos actualmente disponíveis para *Streptomyces* são insuficientes e não reúnem as características de estabilidade e previsibilidade necessárias para a biologia sintética.

Neste trabalho, uma lista de 12 genes estáveis e altamente expressos foi gerada com base em dados de transcriptómica e proteómica publicamente disponíveis para o organismo modelo *Streptomyces coelicolor* A3(2). A caracterização *in silico* das regiões promotoras destes genes permitiu desenhar três promotores minimizados quanto ao tamanho e à complexidade. Os promotores gerados foram caracterizados em *Streptomyces coelicolor* M145 e em *Streptomyces lividans* 1326 ao longo do tempo de cultura, sob influência de choque térmico e em meios de cultura com fontes de carbono definidas. Assim, este trabalho estabelece uma nova estratégia para desenhar promotores para *Streptomyces.*

Palavras-chave: *Streptomyces*, RNA-seq, promotores, análise *in silico*

Table of contents

List of figures

List of tables

List of abbreviations

5'-UTR - 5' Untranslated region

Am - Apramycin

Am^r - Apramycin resistance marker

Ap - Ampicilin

Ap^r - Ampicillin resistance marker

attB - Attachment site (bacteria)

attP - Attachment site (phage)

bp -Base pair

CIRCE - Controlling inverted repeat of chaperone expression

ddH_2O - Double deionized water

DMSO - Dimethyl sulfoxide

DNA - Deoxyribonucleic acid

dNTP - Deoxyribonucleotide

dre - DasR-responsive element

DTT - Dithiothreitol

EDTA - Ethylenediaminetetraacetic acid

GAPDH - Glyceraldehyde 3-phosphate dehydrogenase

gDNA - Genomic deoxyribonucleic acid

GR - GeneRuler

GUS - β-Glucuronidase

kbp - Kilobase pairs

LB - Lysogeny broth

Mbp - Megabase pairs

mRNA - Messenger ribonucleic acid

MS - Mannitol soya flour medium

NMMP - Minimal liquid medium

OD_{600nm} - Optic density at 600 nm

PCR - Polymerase chain reaction

PMSF - Phenylmethanesulfonylfluoride

PNP - p-Nitrophenol

PNPG - p-Nitrophenyl-β-D-glucopyranoside

r.p.m. - Revolutions per min

RBS - Ribosome binding site

RNA - Ribonucleic acid

RNA-seq - RNA sequencing

rRNA - Ribosomal ribonucleic acid

RT-qPCR - Reverse transcription-quantitative polymerase chain reaction

SD -Shine-Dalgarno sequence

SigB - Sigma factor B

Sp - Spectinomycin

SP - Synthetic promoter

SPL - Synthetic promoter library

Sp^r - Spectinomycin resistance marker

TAE - Buffer solution containing Tris base, acetic acid and EDTA

TSB - Tryptone soya broth

TSS - Transcriptional start site

TT - Transcriptional terminator

WT - wild-type

X-Gluc - 5-Bromo-4-chloro-3-indolyl β-D-glucuronide

YEME - Yeast extract-malt extract medium

Introduction

Streptomyces

The genus *Streptomyces* was firstly described by Waksman and Henrici, in 1943 [1]. *Streptomyces*, from the Greek *streptos* (bent) and *mukês* (fungus), refers to a group of filamentous, high GC content, Gram-positive bacteria, harbouring more than 650 species [2]. The life cycle of these bacteria in solid media presents three main stages: the growth of vegetative mycelia, the formation of aerial mycelia and the differentiation of the latter into chains of spores (**figure 1**). *Streptomyces* spp. are chemoorganotrophic and are able to use a wide range of organic compounds, even from complex polymeric substrates. They are widespread in nature and can be found in a wide range of habitats (mostly in the soil and freshwaters) due to the extensive ranges of temperature and pH they tolerate. The secondary metabolism of these bacteria is amongst the most prolific sources of natural bioactive products known [3].

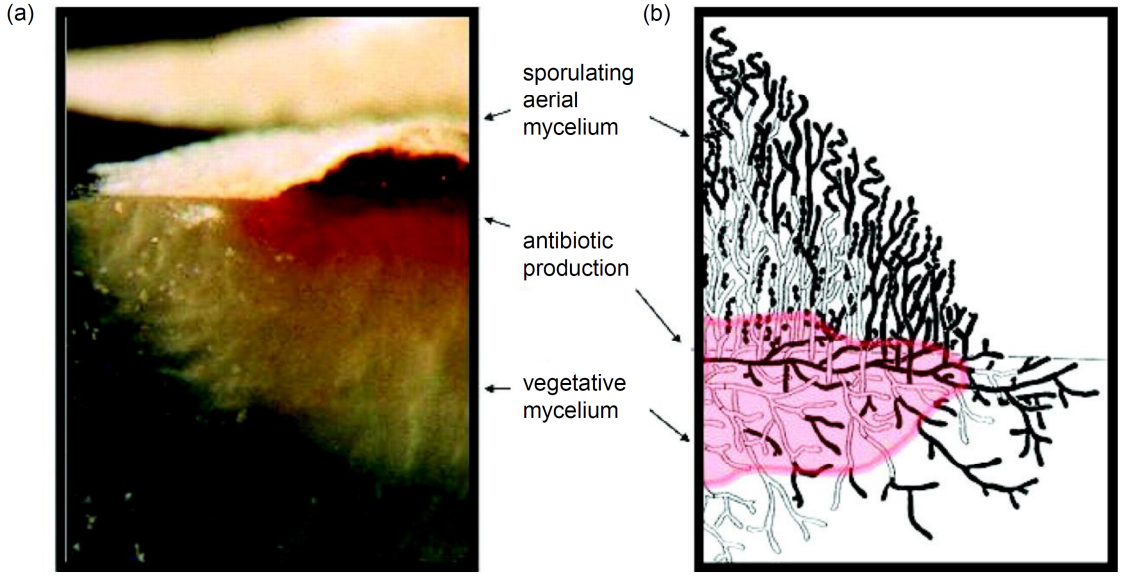


Figure 1. Vertical sections through a *Streptomyces* colony. Photograph (**a**) and scheme (**b**) of colony growing on agar. In the scheme, dead cells are represented in white and living cells in black. Adapted from [4].

Streptomyces coelicolor A3(2) was the first representative of the genus to be fully sequenced and is used in most genetic studies [5]. Its genome is constituted by a 8.7 Mbp linear chromosome, featuring 7825 coding sequences, and two circular plasmids, SCP1 and SCP2, with 356 kbp and 31 kbp, respectively [6-8]. This sequencing project shed light upon the biosynthetic potential of the genus, unveiling a high number of gene clusters dedicated for secondary metabolism. *S. coelicolor* M145, a *S. coelicolor* A3(2) derivative strain lacking the SCP1 and SCP2 plasmids, has been comprehensively

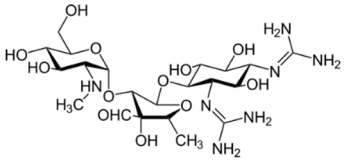
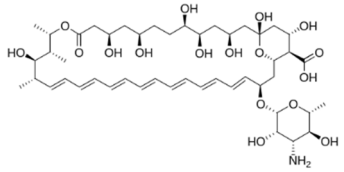
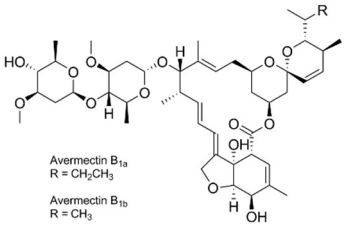
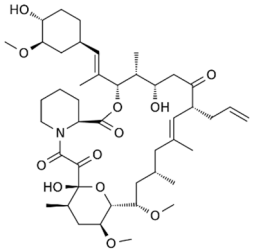
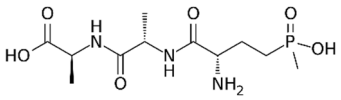
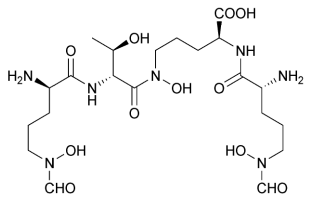
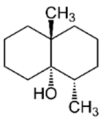
studied, namely regarding its transcriptome and proteome [9-11]. Understanding the genetic expression mechanisms in *Streptomyces* model organism using *-omics* approaches is arguably a major advantage for applied projects involving these organisms, namely in the optimization of the production titters of secondary metabolites.

Secondary metabolism

The genus *Streptomyces* is known to produce over 7600 natural bioactive metabolites and, remarkably, each strain has the genetic potential to produce several compounds [12]. The metabolites are grouped according their function as antagonistic agents, pharmacological agents, agro-biologicals or compounds with regulatory functions [13]. The secondary metabolites found in *Streptomyces* are biochemically derived from the concerted action of enzymes or enzymatic complexes generally encoded in genomic islands as clusters, some of which conserved amongst species [14]. The metabolite classes are related to the biosynthetic pathways and they include polyketides derived from polyketide synthases, peptides from non-ribosomal peptide synthetases, terpenoids, siderophores, aminoglycosides and their combinations and derivatives [14]. A representative panel of metabolites derived from *Streptomyces* is shown in **table 1**.

The number of sequencing projects of *Streptomyces* has been increasing at high rates since 2002, with 19 completely sequenced genomes and 125 draft genome as of May 2014 [15]. The analysis of secondary metabolites and biosynthetic gene clusters underwent great improvements thanks to the computational tools developed to identify these genomic clusters [16]. For example, antiSMASH is a web tool for genome mining of biosynthetic gene clusters that relies on gene organization and protein domain homologies [17]. The genome sequencing and annotation projects have unveil that *Streptomyces* devote up to 10% of their coding capacity to secondary metabolism related genes [18]. On the other hand, most of the detected biosynthetic gene clusters are not associated with previously isolated metabolites. These clusters whose product have not been detected under laboratory conditions are called silent or cryptic clusters. In *S. coelicolor* A3(2), from the 31 secondary metabolite gene clusters identified, only sixteen have been associated with detected metabolites [14]. The awakening of cryptic gene clusters using synthetic biology approaches has been described as a promising technique for the production of novel secondary metabolites [19]. Using the proper strategies, the vast unexplored biosynthetic libraries of *Streptomyces* could reveal themselves as a major source of compounds with utility to mankind.

Table 1. Examples of secondary metabolites produced by *Streptomyces* spp.

 **Streptomycin** [20]	**Function:** Antagonistic agent **Class:** Aminoglycoside **Producer:** *S. griseus* **Observations:** First treatment for tuberculosis
 **Amphotericin B** [21]	**Function:** Antagonistic agent **Class:** Polyketide **Producer:** *S. nodosus* **Observations:** Treatment of cryptococcal meningitis and visceral leishmaniasis
 **Avermectins** [22]	**Function:** Antagonistic and agrobiological agent **Class:** Polyketide **Producer:** *S. avermitilis* **Observations:** Treatment for parasitic worms, insecticide against lice and mites
 **Tacrolimus** [23]	**Function:** Pharmacological agent **Class:** Hybrid polyketide and non-ribosomal peptide **Producer:** *S. tsukubaensis* **Observations:** Immunosuppression following organ transplantation
 **Bialaphos** [24]	**Function:** Agrobiological **Class:** Non-ribosomal peptide **Producer:** *S. hygroscopicus* **Observations:** Broad-spectrum herbicide, transgenic plant selection
 **Coelichelin** [25]	**Function:** Regulatory functions **Class:** Non-ribosomal peptide **Producer:** *S. coelicolor* A3(2) **Observations:** Iron chelator
 **Geosmin** [26]	**Function:** Regulatory functions **Class:** Terpenoid **Producer:** Several *Streptomyces* spp. **Observations:** Earthy odour

Synthetic biology in *Streptomyces*

Synthetic biology has been defined as the engineering of biology through the synthesis of complex, biologically based (or inspired) systems, which display functions that do not exist in nature [27]. Synthetic biology aims to design and construct new genetic circuits and to remake natural existing ones for useful purposes. A key idea of synthetic biology approaches is the plug-and-play genetic manipulation (**figure 2**) [28]. The fact that secondary metabolite biosynthetic genes are already organized in clusters simplifies the application of such idea. A strategy for the activation of the cryptic clusters would consist on the complete removal of the native regulation and the substitution by a synthetic, predictable regulation. Once the strategy had been optimized, it could function as scaffold for high-throughput screening of the vast number of annotated clusters encoding for the same class of compounds [29]. Besides the awakening, the manipulation of synthetic assemblies would allow to exploit the modular nature of secondary metabolism enzymatic complexes to produce novel or optimized derivatives from a given biosynthetic pathway. Synthetic biology approaches rely on fully characterized genetic components, standardized assembly procedures and optimized hosts, also known as chassis, to achieve predictable behaviours and products in living systems. The synthetic biology toolbox for *Streptomyces* is based in a long tradition of genetic studies and a recent focus on strain enhancement and toolbox development *per se,* that will be reviewed in the following sections.

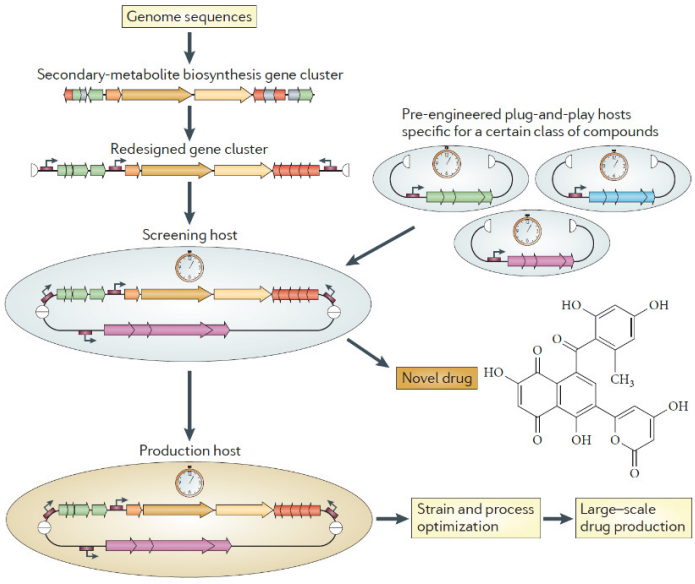


Figure 2. Plug-and-play strategy for heterologous expression of cryptic biosynthetic pathways. The redesign of gene clusters of secondary-metabolite biosynthetic proceeds from the streamlining of the coding sequences with inclusion of known transcriptional elements (represented by small arrows). Reproduced from [28].

Streptomyces-*based chassis*

The natural ability of *Streptomyces* bacteria of producing secondary metabolites fits them into the optimal choice for development of chassis for the heterologous expression of natural or engineered biosynthetic gene clusters. The strains chosen should fulfil the following characteristics: high growth rate, genomic stability, compatible molecular tools and a surplus of biosynthetic precursors. There are two main groups of *Streptomyces*-based chassis: wild-type strains and engineered strains. The wild-type strains used as heterologous hosts have their genome sequenced and are genetically amenable to work with. *S. lividans* 66 (or 1326) is an example of such wild-type strain: it was sequenced and it is naturally tolerant to exogenous methylated DNA [30]. *S. albus* J1074, a *Sal*I-defective derivative of *S. albus* G with naturally-minimized genome [31], presented improved yields in secondary metabolite production relatively to *S. lividans* [32]. The engineered strains are generally derived from extensively known species by rational genome minimization, including deletion of unstable or precursor-diverting genomic regions. Several *S. coelicolor* A3(2) derivatives were produced by deletion of the gene clusters encoding the detectable metabolites actinorhodin, calcium-dependent antibiotic, prodiginine and yCPK and/or by modifications in the transcriptional and translational machinery [33]. Mutants of *S. avermilitis* with deletion of more than 1.4 Mbp that include the biosynthetic cluster for avermectins, oligomycins and filipins were also generated and characterized [34]. These optimized strains have been presented as the most adequate chassis to accommodate exogenous genetic circuits in order to achieve the expected outputs [35].

Genome editing techniques

The development of molecular tools and techniques is an important requirement for strain manipulation. The delivery of exogenous DNA to *Streptomyces* is mainly achieved by intergeneric conjugation [36, 37], a highly effective technique based in the horizontal gene transfer of an oriT-containing plasmid from a suitable *tra*-encoding *E. coli* host to *Streptomyces*. Genetic manipulation procedures in *Streptomyces* chromosome are based in homologous recombination events. One of the most widely used techniques is REDIRECT [38], a PCR-targeting and λ-Red mediated recombination methodology where the target sequence is replaced by cassette containing a selectable antibiotic resistance flanked by the yeast FLP-recombinase target sequences for selective marker removal. The deletion of large genomic regions has been achieved most of the times by *Cre-loxP* system from the P1 phage [39]. Homologous recombination is used for introduction of *loxP* sites flanking the target region followed by *Cre* expression for excision of flanked site. Recently, the Cas9/CRISPR technology,

based in a primitive bacterial immune system [40], was proven its utility in *Streptomyces* genome editing, including large deletion and point mutations [41, 42]. This technique is based in homologous recombination-based repair following a RNA-guided cleavage of the genomic target by Cas9 nuclease. Methodologies for introduction of exogenous DNA into *Streptomyces* include use of self-replicating and integrating vectors. For example, the SCP2* derivatives are self-replicating plasmids that presents a low copy number, stable inherence and being capable of carrying >30 kbp inserts [43]. The integrative vectors are based in the genomic integration systems of *Φ*C31 and *Φ*BT1 actinophages into the respective bacterial attachment sites (*attB)* encoded in several *Streptomyces* spp., including *S. coelicolor* and *S. lividans* [44, 45]. These integrative vectors contain a phage-derived integrase and the respective attachment site (*attP)* which allows, in theory, a single site-specific unidirectional recombination. A vast array of integrative vectors have been created, from expression [46] and reporter vectors [47, 48] to high capacity vectors such as bacterial artificial chromosomes [49] and cosmids [44].

Molecular toolbox for Streptomyces

A molecular toolbox is constituted by genetic parts that suit defined functions in transcriptional and translational control or gene encoding. The assembly of these parts allows to create molecular devices with defined purposes. The implementation of such devices in suitable chassis yields the desired output. Thus, the genetic parts, represented by promoters, ribosome binding sequences, coding sequences and transcriptional terminators represent the basis for synthetic biology approaches. The rational development of new biological devices rely on comprehensive characterization of the genomic parts employed [50].

In bacteria, transcription is driven by RNA polymerase complex, constituted by the elongation-capable core with subunits $\alpha_2\beta\beta'\omega$ and the initiation-required sigma factor. The DNA sequences that are able to recruit a sigma-containing RNA polymerase are called promoters. The most used promoter for heterologous gene expression in *Streptomyces* has been *ermE*p*, a derivative of erythromycin resistance gene promoter of *Saccharopolyspora erythraea* (formerly, *Streptomyces erythraeus*) [51, 52]. Some promoters are under influence of transcriptional regulators. The regulation determinants are the operator, usually a conserved DNA motif in the promoter region, and the regulator, a protein that can recognize and bind the operator. A repressor is a regulator that inhibits transcription by binding its cognate operator in the absence of an inducer signal. Thereby, promoters regulated by repressors are dependent on the presence of the inducer for transcriptional activity and are known as inducible promoters. Most inducible promoters used as molecular tools for *Streptomyces* are listed in **table 2**. While

some of the reported inducible systems were based in naturally regulated promoters, others were synthetically derived from operator-free promoters merged with known operator/repressor systems.

Table 2. Inducible promoters based in repressors used in *Streptomyces*.

Promoter	Repressor	Inducer	Promoter source	References
*gyl*p1/p2	GylR	Glycerol	*S. coelicolor* A3(2)	[53]
*nitA*p	NitR	ε-caprolactam	*Rhodococcus rhodochrous* J1	[54]
tcp830	TetR	Tetracycline Anhydrotetracycline	Synthetic	[55]
*cpkO*p	ScbR	γ-butyrolactones	*S. coelicolor* A3(2)	[56]
P21-cmt	CmtR	Cumate	Synthetic	[57]
PA3-rolO	RolR	Resorcinol	Synthetic	

Activators are regulators that positively affect the transcriptional activity of a promoter in the presence of an inducer. The *tipA*p is a popular activator-based inducible promoter for *Streptomyces*. In the presence of thiostrepton, TipA is overexpressed and recruited to *tipA* promoter, activating its own transcription [58-60]. Exogenous promoter/RNA polymerase systems, based on the T7 phage machinery, were adapted for *Streptomyces*, allowing fully orthogonal transcriptional regulation [61].

The assessment of promoter activity can be performed by cloning the query sequence upstream of reporter genes. The reporter genes generally code for enzymes which activity can be measured with high sensitivity and specificity. The specificity is a major issue in reporter gene development for *Streptomyces* given its natural richness in catabolic enzymes and antibiotic production (and consequently, resistance). The reporter genes described for *Streptomyces* are summarized in **table 3**.

Translation is a further layer of gene expression regulation. The 70S ribosomes are the molecular machines responsible for translation in bacteria, being constituted by two asymmetric rRNA-protein complexes, the 30S and 50S subunits. Interestingly, several antibiotics produced by *Streptomyces* spp. target ribosomes [62].

Table 3. Reporter genes used in *Streptomyces*. MIC, minimal inhibitory concentration.

Reporter gene	Protein encoded	Measurements	Source	References
neo	Aminoglycoside phosphotransferase	MIC Enzymatic assays	Tn5 transposon	[63]
cat	Chloramphenicol O-acetyltransferase	MIC Enzymatic assays	Tn9 transposon	[64]
aacC1	Aminoglycoside O-acetyltransferase	MIC Enzymatic assays	Tn1696 transposon	[65]
melC	Tyrosinase	Enzymatic assays	*S. glaucescens*	[66]
luxAB	Luciferase	Bioluminescence	*Photorhabdus luminescens*	[67]
gfp	Green fluorescent protein	Intrinsic fluorescence	*Aequorea victoria*	[68, 69]
xylE	Catechol 2,3-dioxygenase	Enzymatic assays	*Pseudomonas putida*	[48]
gusA	β-glucuronidase	Enzymatic assays	*E. coli*	[47]
bpsA	Indogoidine synthase	Indogoidine quantification	*S. aureofaciens* *S. lavendulae*	[70, 71]

Although exciting breakthroughs in ribosome engineering have been achieved in other organisms [72], translation control in *Streptomyces* is mainly attained by modulating the affinity of mRNA towards ribosomes. In bacteria, the ribosome binding site (RBS, also known as Shine-Dalgarno [SD] sequence) is a sequence in the 5' untranslated region (5'-UTR) of mRNAs that present some complementarity with the 3' region of the 16S rRNA from 30S subunit and is involved in mRNA-ribosome recognition and positioning. The usage of RBS from genes with the desired expression level is the most common technique for tuning translational efficiency in *Streptomyces* [29]. Translation efficiency can also be affected by secondary structures of mRNA. Riboswitches are genetic control elements present in 5'-UTRs of mRNA that can generate alternative secondary structures through binding of a given metabolite, such as the B12 riboswitch in *S. coelicolor* A3(2) [73]. Synthetic riboswitches for *Streptomyces* have been designed using theophylline-sensitive aptamers that conditionally allow the access of ribosome to RBS and start codon of mRNAs, thereby exerting ligand-dependent translational control (**figure 3**) [74].

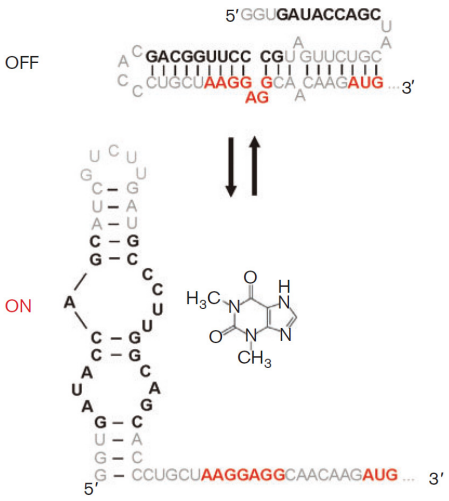


Figure 3. Alternative structures of the theophylline-sensitive E* riboswitch. The SD sequence, represented in red, is made accessible in the presence of theophylline. Reproduced from [74].

The advances in chassis, techniques and molecular tools for synthetic biology approaches in *Streptomyces*, although promising, are yet very modest compared with other organisms. The complex transcriptional apparatus and dynamics in *Streptomyces* still hampers the development of novel parts such as constitutive promoters.

Development of constitutive promoters for *Streptomyces*

In bacteria, transcription of DNA into mRNA by RNA polymerase is the major checkpoint for controlling gene expression. For controlling transcription, there are two main factors to be considered: the existence of regulators bound to operators in promoter regions and the sigma factors available for promoter recognition. In *Streptomyces*, each of these factors is represented by a massive number of variables: for the model organism *S. coelicolor* A3(2) a total of 499 transcriptional regulators and 64 sigma factor have been reported [75]. A successful strategy for transcriptional control in *Streptomyces* would need to circumvent these complex native networks. Three scenarios could be considered: (i) to implement a totally exogenous transcriptional machinery, such as the T7 RNA polymerase system [61], (ii) to rely on exogenous orthogonal regulator/operator systems [76] or (iii) to remove native regulation of stably expressed components in order to turn them into predictable parts. In the context of constitutive promoters' development for *Streptomyces*, the latter strategy has been applied in a more or less declared way. Predictable promoters are indispensable tools for gene cluster refactoring. Furthermore, well-characterized promoters are building blocks for synthetic inducible systems and can be combined with riboswitches, providing a myriad of new components for synthetic biology approaches.

Determinants of constitutive promoters in Streptomyces

Constitutive promoters are those active at all circumstances, which, in practice, is understood as being recognized by the primary sigma factor without need of other transcriptional factors [77]. The core promoter is the minimal portion of a promoter region required to initiate transcription and it is constituted by two sigma-recognizable hexamers centred approximately 10 and 35 bp upstream of the transcription start site (TSS) and the spacing between them. The primary sigma factor in *S. coelicolor*, HrdB [78, 79], belongs to the sigma-70 family and recognizes the consensus sequence depicted in **figure 4** [80].

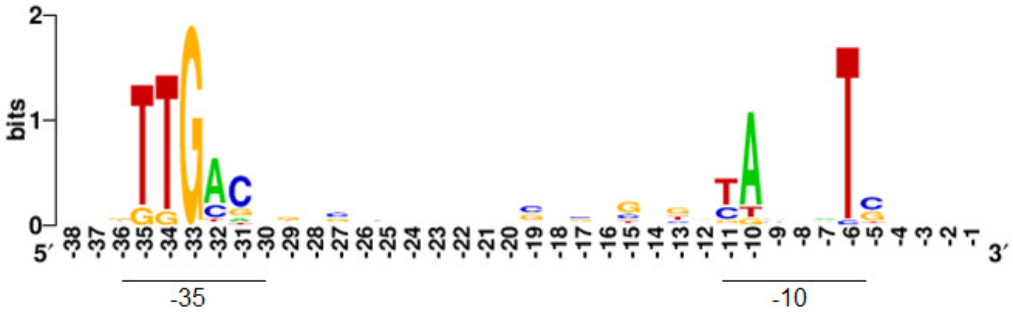


Figure 4. Representation of the sequence recognized by HrdB in *Streptomyces*. The Weblogo representation was derived from 29 *Streptomyces'* promoters resembling sequences recognized by sigma-70 in *E. coli* [80].

The promoters of *Streptomyces* were classified based on clustered frequency and positional analysis of over-represented short sequences, including motifs predominantly centred in the -35 and -10 position relatively to the TSS [81]. This analysis revealed the prevalence of the sigma-70-compatible motifs TANNNT for -10 region and TTGAC for -35 region, but also the extended motifs TNTNNNANNT, TGNNANNNT and GTNNANNNT centred at -10 region, similar to what occurs in *E. coli* and *Corynebacterium glutamicum* [82, 83]. Recently, a RNA-seq study revealed an extended motif centred in -35 region of promoters of transcriptional machinery of *S. coelicolor* A3(2) (**figure 5**) [9].

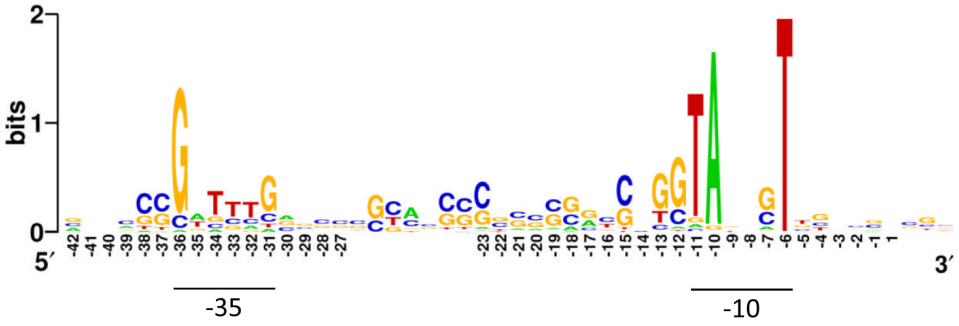


Figure 5. Representation of conserved sequence in promoters associated with transcriptional machinery in *Streptomyces*. Adapted from [9].

Strategies for design of constitutive promoters for Streptomyces

The first and most used "synthetic" constitutive promoter for *Streptomyces* is actually the result of a point deletion in the erythromycin resistance gene (*ermE*) promoter of *Saccharopolyspora erythraea* [52]. The genomic region upstream to *ermE* harbours a complex web of RNA polymerase-attracting sites, namely, two promoter in the forward strand, *ermE*p2 and *ermE*p1, and a promoter in the reverse strand [51]. Together, *ermE*p2 and *ermE*p1 constitute *ermE*p (**figure 6**). The TGG deletion from the -35 region of *ermE*p1 originated a 1.08-fold stronger promoter, known as *ermE*p* [52]. This effect was suggested to result from the more HrdB-complying -35 region. The -10 region of e*rmE*p1 was identified as belonging to the -10 extended class, TGNNANNNT.

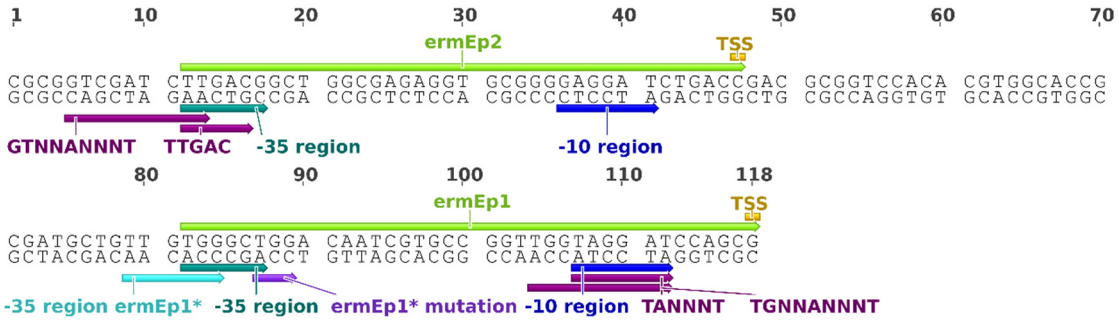


Figure 6. Representation of the *ermE* promoter. The *ermE*p1 and *ermE*p2 promoters are depicted by light green arrows. The motifs associated with core sequences of *Streptomyces*' promoters [81] are depicted by dark purple arrows. The remaining annotations are derived from the original characterizations [51, 52].

In 1997, Labes *et al.* described SF14 [65], a promoter-containing fragment derived from the *S. ghanaensis* phage I19 responsible for 1.66-fold higher kanamycin resistance levels than *ermE*p*. SF14 is also constituted by two tandem promoters, 14-Ip and 14-IIp, but in contrast with *ermE*p, their -10 regions partially overlap (**figure 7**). One TTGAC motif and one TANNNT are present in the -35 and -10 regions of 14-IIp, respectively. Both 14-Ip and 14-IIp promoters include the HrdB-recognizable sequence and were validated by *in vitro* run-off transcription assays using fractions containing HrdB-enriched RNA polymerase holoenzyme.

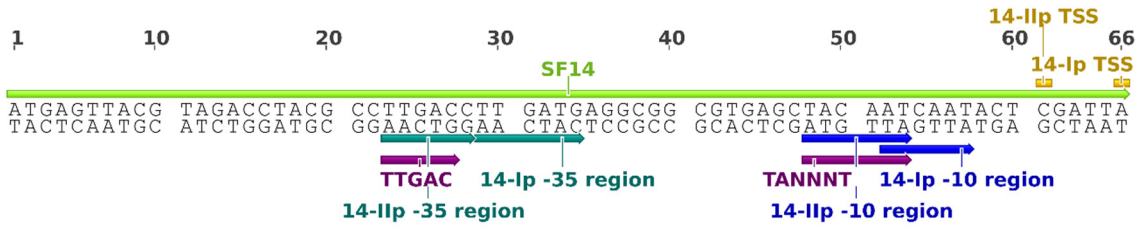


Figure 7. Representation of the SF14 promoter. The motifs associated with core sequences of *Streptomyces'* promoters [81] are depicted by dark purple arrows. The remaining annotations are from the original characterization [65].

*kasO*p* was generated by deleting the ScbR operators present in the inducible *cpkO* promoter [84]. *kasO*p* was recognized by HrdB *in vivo* and presented a constitutive pattern of transcription. Using *neo* as reporter gene, *kasO*p* presented 2-fold higher promoter activity than *ermE*p* in *S. coelicolor* A3(2). The presence of a single HrdB-binding site with 18-nt spacer was regarded as avoiding steric hindrances between RNA polymerases. The TANNNT and TTGAC motifs are present in the -10 and -35 regions of *kasO*p*. Besides, three additional -10 motifs are represented in its sequence (**figure 8**).

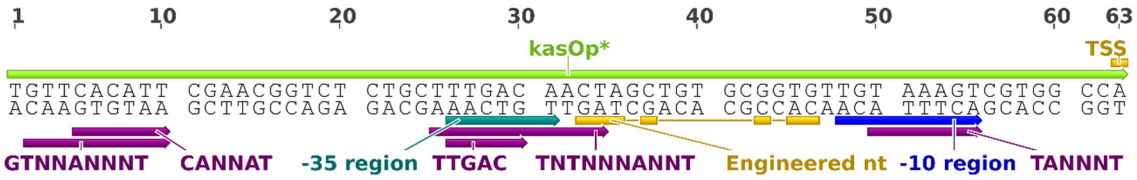


Figure 8. Representation of kasOp*. The motifs associated with core sequences of *Streptomyces'* promoters [81] are depicted by dark purple arrows. The remaining annotations are from the original characterization [84]. The *engineered nucleotides* annotation refer to the mutations in ScbR operator.

The hint for synthetic promoter libraries (SPL) construction came after the knowledge that by keeping the -35 and -10 regions and randomizing the spacer sequence one could modulate the strength of prokaryotic promoters [85]. The first attempt to produce SPL for *Streptomyces* spp. was based in the -35 and -10 hexamers of HrdB-recognizable sequence, using the degenerate oligonucleotide N_{10}TTGACNN$_{17}$TASVDTN$_5$ [86]. Noteworthy that the -10 region was also partially randomized according to the consensus used, with 18 possible sequences. The strongest representative of the library, A1-14, displayed 0.92-fold activity when compared to *ermE*p*. A1-14 contains the TTGAC and TANNNT motifs, which were ubiquitous in the library, but also a -10 extended class motif, TGNNANNNT (**figure 9**).

Clustered analysis of strong versus weak synthetic promoters of the SPL showed that: (i) TAGGGT would be the typical -10 region of strong promoters, (ii) the motif RGgGn immediately upstream -10 region is an extension present only in strong promoters and (ii) that a G-rich spacer would be influential in strength of streptomycete promoters. Furthermore, it was suggested that imperfect repetitions of -10 region might help RNA polymerase positioning.

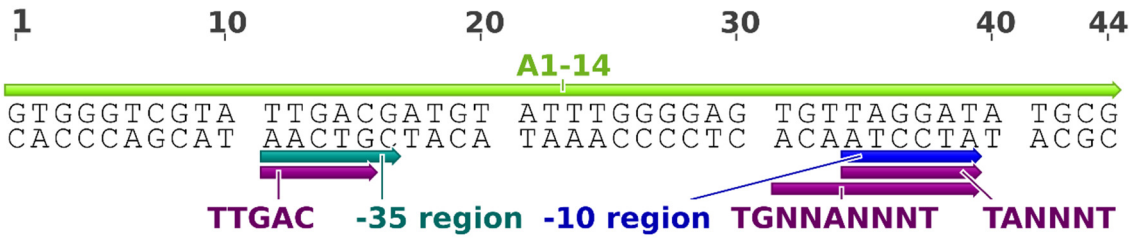


Figure 9. Representation of the A1-14 promoter, the strongest representative of a SPL based in the sequence recognized by HrdB. The motifs associated with core sequences of *Streptomyces'* promoters [81] are depicted by dark purple arrows. The remaining annotations are from the original characterization [86].

In 2013, Siegl *et al.* reported and characterized a new SPL based in the -10 and -35 regions of *ermE*p1, $N_6GGCTN_{19}TAGGATN_6$ [87]. P21 promoter (**figure 10**) displayed 1.67- and 2.92- fold higher activity, in *S. lividans* TK24, than *ermE*p* and *ermE*p1, respectively. The promoter showed similar results in several different actinomycetes. By RNA-seq analysis, the authors showed that the TSS of P21 promoter is located 6 nucleotides downstream the -10 region.

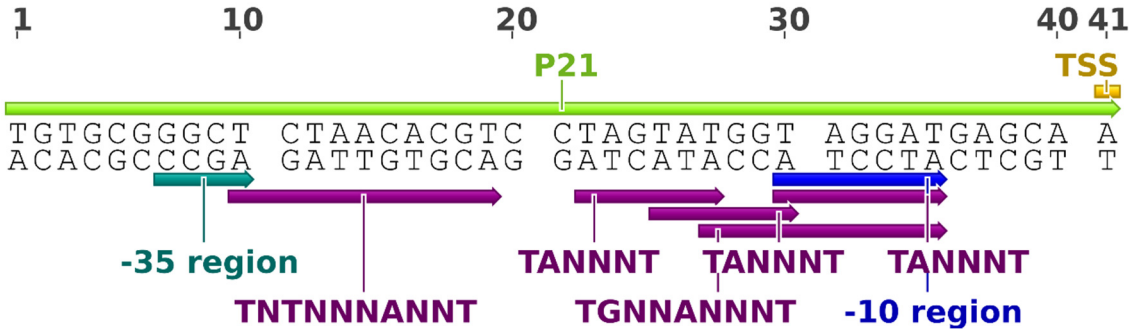


Figure 10. Representation of the P21 promoter, the strongest representative of a SPL based in *ermE*p1's -10 and -35 regions. The motifs associated with core sequences of *Streptomyces'* promoters [81] are depicted by dark purple arrows. The remaining annotations are from the original characterization [87].

Several randomized nucleotides are conserved between *ermE*p1 and P21, including the TGG motif immediately upstream the -10 region, which classifies this region as extended (motif TGNNANNNT). Three TANNNT elements - TAGTAT, TATGGT and TAGGAT - are present in P21 promoter, surrounding the -10 region. The -35 region consensus observed in P21 does not comply with the usual TTGAC motif, albeit the G, the most conserved nucleotide of the motif, is present.

Looking for a set of constitutive promoters for gene cluster reengineering, Shao *et al.* [88] characterized the promoter activity of the genomic regions upstream of the coding sequences for glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and 30S ribosomal protein S12 (rpS12, **figure 11**) of *S. griseus,* which they had found highly transcribed along time by RT-qPCR among 23 housekeeping genes. These genomic

regions promoted over 10-fold higher reporter activities than *ermE*p*, in *S. lividans* 1326. A whole set of 13 intergenic regions presenting promoter activity was completed from the sequences upstream *gap1* and *rpS12* of other actinobacteria. The cloned regions were about 300-nucleotide long and were putatively assumed to have RBSs in 6-10 bp upstream start codon, based on the presence of a AG-rich region.

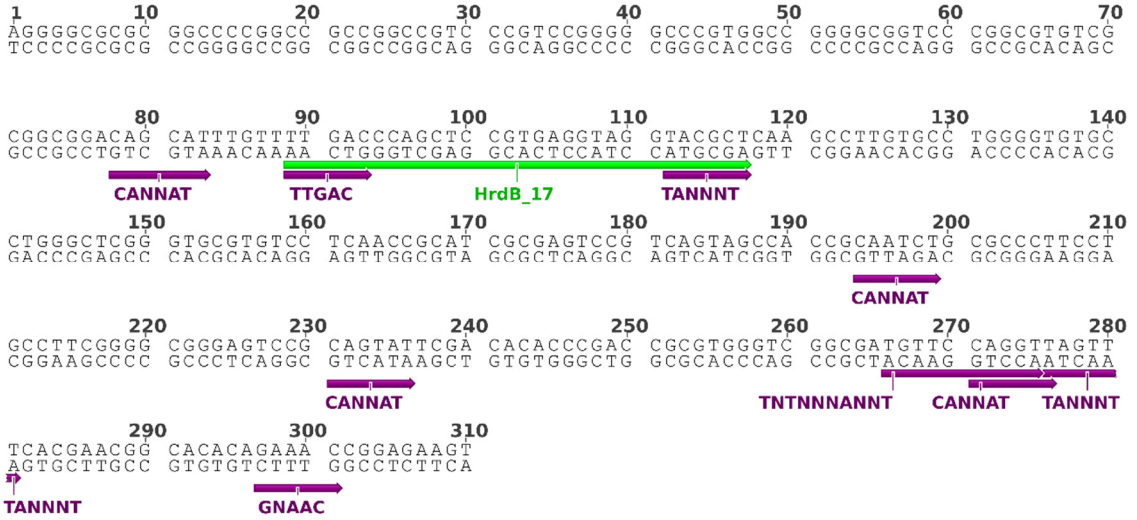


Figure 11. Representation of the genomic region upstream of *rpS12* in *S griseus*. The promoter activity of this region was assessed by Shao *et al* [88]. The motifs associated with core sequences of *Streptomyces'* promoters [81] are depicted by dark purple arrows. The region matching the sequence recognized by HrdB with a spacer of 17 nucleotides and evidence of promoter core sequences is depicted by the green arrow.

The design of constitutive promoters for *Streptomyces* followed a path from the usage of naturally occurring, complex genomic regions of related organisms, to rational trimming and promoter engineering of known inducible promoters, coming to consensus-based randomization of spacers and synthetic promoter library construction. During the course of these developments, some determinants of strong constitutive promoters were added to the consensus and motifs found during the first studies in the early '90s. The last reported attempt to obtained strong promoters [88] reintroduced the idea of using native promoters, now selected by previous low-scale transcriptional studies. Indeed, several never-so-strong promoters were reported using this approach. However, the rising rationality in promoter choice/generation was withdrew by using full genomic regions, which could contain unexpected operators or encode 5'-UTR that acquire secondary structures that hamper translation. Hence, the combination of the increasing whole-genome transcriptomic data, mainly available for the model organism *S. coelicolor* A3(2), with the rational principles of trimming, operator freeing and compliance with strong promoters' motifs, seemed as the next logical strategy to design constitutive promoters for *Streptomyces.*

Objectives

The current molecular toolbox for *Streptomyces* includes several useful representatives of constitutive promoters, such as *ermE*p* and SF14. However, the limited number of promoters together with the complexity and unpredictability of the existing ones still hinders genetic engineering progresses. The increasing number of whole genome transcriptome/proteomic studies and bioinformatics developments pave the way for new rational approaches to design synthetic parts. The present work aims to apply these assets to establish novel constitutive promoters for *Streptomyces.* The objectives of this work were:

- Selection of stable and highly expressed genes of *S. coelicolor* M145 using publically available whole-genome transcriptomic and proteomic data;
- *In silico* characterization of genomic region upstream of the selected genes regarding functional sequences, such as transcriptional terminators, core promoters and ribosome binding sites;
- Design of synthetic promoter candidates according to the described genomic regions of *S. coelicolor* M145;
- *In vivo* characterization of the synthetic promoters in the model organism *S. coelicolor* M145 and in a common chassis for heterologous gene expression, *S. lividans* 1326, using promoter probe vectors.

Material and methods

Bacterial strains, plasmids and growth conditions

All strains and plasmids used in this work are listed in **table 4**. *Streptomyces coelicolor* M145, *S. lividans* 1326 and their derivative mutant strains were grown at 28 ºC, at 220 rpm. For DNA extraction procedures, *Streptomyces* spp. were grown in YEME medium [89]. For synthetic promoter strength determination, bacteria were grown in TSB [89] or NMMP medium [89, 90]. Sporulation was achieved in MS agar medium, at 28 ºC. Spores of *S. coelicolor* and *S. lividans* were collected using a solution of 30% (w/v) glycerol and 0.021% (v/v) Triton X-100, spectrophotometrically quantified using an UVmini-1240 UV-Vis Spectrophotometer (Shimadzu) and kept at -80 ºC. For the determination of cells dry weight, 1-mL aliquots were harvested, washed once with 0.9% (w/v) NaCl solution and dried at 80 ºC for at least 48 h. *E. coli* was routinely grown in Lysogeny broth (LB), at 37 ºC. Stock cultures of *E. coli* were kept in 33% (w/v) glycerol at -80 ºC. When necessary, medium was supplemented with antibiotics to final concentrations of 50 μg.mL^{-1} apramycin (Am), 100 μg.mL^{-1} ampicillin (Ap), 50 μg.mL^{-1} kanamycin, 25 μg.mL^{-1} chloramphenicol, 50 μg.mL^{-1} spectinomycin (Sp) and 25 μg.mL^{-1} nalidixic acid**.**

Table 4. Strain and plasmids used in this work.

Strain or plasmid	Main characteristics	Reference
Streptomyces		
S. coelicolor M145	SCP1$^-$ SCP2$^-$ derivative from A3(2) strain	[89]
S. lividans 1326	wild-type	[30]
Escherichia coli		
DH5α	General cloning strain [F$^-$ φ80*lacZ*ΔM15 Δ(*lac*ZYA-*arg*F)U169 *recA*1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 thi$^-$1 *gyr*A96 *rel*A1 λ$^-$]	[91]
ET12567 [pUZ8002]	Non-methylating strain for intergeneric conjugation [*dam dcm hsdM hsdS hsdR* Tc^r Cm^r] [RP4 plasmid derivative, Km^r *tra* operon]	[38, 92]
Plasmids		
pUC19	General cloning vector, Ap^r *lacZα* pMB1ori	[93]
pGUS	Promoter probe vector for *Streptomyces* [*ΦC31-integrase gusA* Am^r Sp^r oriT pMB1ori *attB*]	[47]
pGUS_SP	pGUS derivative with synthetic promoter fragment cloned into XbaI and KpnI sites of pGUS	This work

Selection and characterization of *S. coelicolor* promoters

The search for stable and highly expressed genes from *S. coelicolor* was based on publically available transcriptomic and proteomic data. The genes with a <5% variation of *read counts per transcript* were retrieved from a time-lapse whole genome transcriptome analysis study of *S. coelicolor* M145 [94] and ranked according to the average transcription level. The list of retrieved genes was conferred with the list of the 65 most abundant proteins retrieved from a semi-quantitative proteomics study in *S. coelicolor* M145 [11]. A final list of *S. coelicolor* genes was obtained that displayed the common entries from transcriptomic and proteomic data and was sorted according a score given by the product of *read counts per transcript* and the *total exponentially modified protein abundance index*. The transcription start site for each gene was annotated according to RNA-seq studies in *S. coelicolor* M145 [9, 10]. The regions upstream of the selected genes were annotated regarding ribosome binding sequences [80], promoter core motifs in *Streptomyces* [81], sequences recognized by HrdB [80], transcriptional terminators and regulatory motifs/operators described in the literature.

In silico analysis of DNA sequences

The annotation of genomic sequences was performed using Geneious version 4.8.3 [95]**.** The intrinsic terminators of transcription in *S. coelicolor* were obtained from WebGesTer Database [96]. The reconstructed operators/motifs were represented using WebLogo [97]. The search of motifs in genomic sequences was done using FIMO [98].

DNA manipulation

Plasmids were isolated from 5 mL cultures of *E. coli* grown overnight in LB medium, at 37 ºC, using the GenElute™ Plasmid Miniprep Kit (Sigma), following the manufacturer's instructions. *Streptomyces* genomic DNA (gDNA) was isolated from 1 mL aliquots of stationary phase cultures using the GeneJET Genomic DNA Purification Kit (Thermo Scientific) according to the manufacturer's instructions. DNA digestion products and PCR products were purified using NZYGelpure Kit (NZYTech) according to the manufacturer's instructions. DNA was eluted with ddH_2O, quantified using NanoDrop™ 1000 Spectrophotometer (Thermo Scientific) and analysed by gel electrophoresis.

PCR conditions

Generic conditions of polymerase chain reactions (PCR) were as follows: 1x GoTaq® Reaction Buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs (Thermo Scientific), 0.2 µM forward primer, 0.2 µM reverse primer, 1.25 U GoTaq™ G2 flexi DNA Polymerase (Promega) and 5% (v/v) of dimethyl sulfoxide (DMSO). All the oligonucleotides used in this work are listed in **table 5**. Template DNA varied between 20 to 50 ng for plasmid DNA and 50 to 100 ng for genomic DNA. All programs included an initial denaturation step (95 ºC, 5 min), a variable number of cycles of denaturation (95 ºC, 30 s), annealing (30 s) and extension (72 ºC, 1 min/kbp) and a final elongation step (72 ºC, 7 min). The reactions were performed in a C1000™ Thermal Cycler (Bio-Rad).

Table 5. Oligonucleotides used in this work.

Name	Sequence (5'-3')	Uses
bla_R	TATATCTAGAGTCTGACGCTCAGTGG	Synthetic promoter generation
P1_F	GGGGTACCTGCTGCTGACGCTACACGGCATGTCCGAGCC TCACCAGTGAGTAAGGGGTGTGCGGAACCCCTATTTG	
P2_F	GGGGTACCCACCTCCGACCCTACCTCTCCGGGGCCTCGG GGGTGACATCGAGACGCCCCGCGGAACCCCTATTTG	
P3_F	GGGGTACCAGGTCCGAGACTATGACTGCGATTAGCACTC GGTCAAGCGGAGGCGGAACCCCTATTTG	
P21_F	GGGGTACCTTGCTCATCCTACCATACTAGGACGTGTTAGA GCCCGCACAGCGGAACCCCTATTTG	
GUSseq_R3	GGTTTCGACGGGCCG	Synthetic promoter sequencing
GUSseq_F2	CGAAGATACCTGCAAGAA	Synthetic promoter presence confirmation
pUCR	ACACAGGAAACAGCTATGAC	
C31_int_F2	TGGGTGTCGCCGTTGGTG	*Φ*C31 *attB* integration confirmation
C31_int_R2	CGTCGTCGGTCGGCGGCT	

Gel electrophoresis

For routine diagnosis, DNA samples were mixed with 1x loading dye and separated by gel electrophoresis (0.8 - 1.2% agarose, 0.2 µg/mL ethidium bromide, 80 - 100 V, 1x TAE [40 mM Tris, 20 mM acetic acid, 1 mM EDTA]). The gels were scanned in a Molecular Imager® GelDoc™ XR+ Imaging System (Bio-Rad). The Lambda DNA/HindIII Marker or GeneRuler™ DNA Ladder Mix (Thermo Scientific) were used as reference for DNA sizing and relative quantification.

Generation and cloning of synthetic promoters

Synthetic promoters P1, P2, P3 and P21 were generated by PCR using the oligonucleotides listed in Table. Briefly, the synthetic promoter was amplified together with the gene coding for beta-lactamase; the 1.2-kbp PCR product was cloned into pGUS plasmid digested with XbaI and KpnI (Thermo Scientific), yielding the pGUS_SP plasmids (pGUS_P1, pGUS_P2, pGUS_P3 and pGUS_P21) (**figure 12**). All constructions were confirmed by DNA sequencing (STAB VIDA, Portugal).

E. coli transformation

For routine cloning procedures, *E. coli* DH5α chemically competent cells were prepared by the divalent cations treatment method [99]. For DNA transformation, 100 µL of cell suspension were mixed with up to 1 µg plasmid DNA and incubated on ice for 20 min. The cells were then heat-shocked at 42 ºC for 45 s and transferred to ice for 2 minutes. The cells were added 900 µL of LB medium and incubated at 37 ºC for 45 min for cell recovery prior to plating in solid medium. *E. coli* ET12567 carrying pUZ8002 plasmid electro-competent cells were transformed with the mobilizable plasmids for *Streptomyces* (**figure 12**) [89]. *E. coli* ET12567 carrying pUZ8002 plasmid was grown overnight in 5 mL LB supplemented with kanamycin and chloramphenicol. 20 mL of fresh medium supplemented with the appropriate antibiotics and 20 mM $MgSO_4$ was inoculated with 200 µL pre-inoculum. The culture was grown at 37 ºC to an OD_{600nm} of 0.4, washed once with 1 volume of ddH_2O and washed twice with 1 volume of 10% (w/v) glycerol. The cells were resuspended in 300 µL of 10% (w/v) glycerol. 50 µL of cells were mixed with 100 ng of pGUS_SP plasmids using a Gene Pulser™ plus Pulse Controller (Bio-Rad) (0.1 cm gap width, 200 Ω, 25 µF, 2.5 kV). The cells were then let to recover in 1 mL LB at 37 ºC for one hour. The cell suspensions were plated in LB agar plates supplemented with chloramphenicol, kanamycin and apramycin and incubated overnight at 37 ºC.

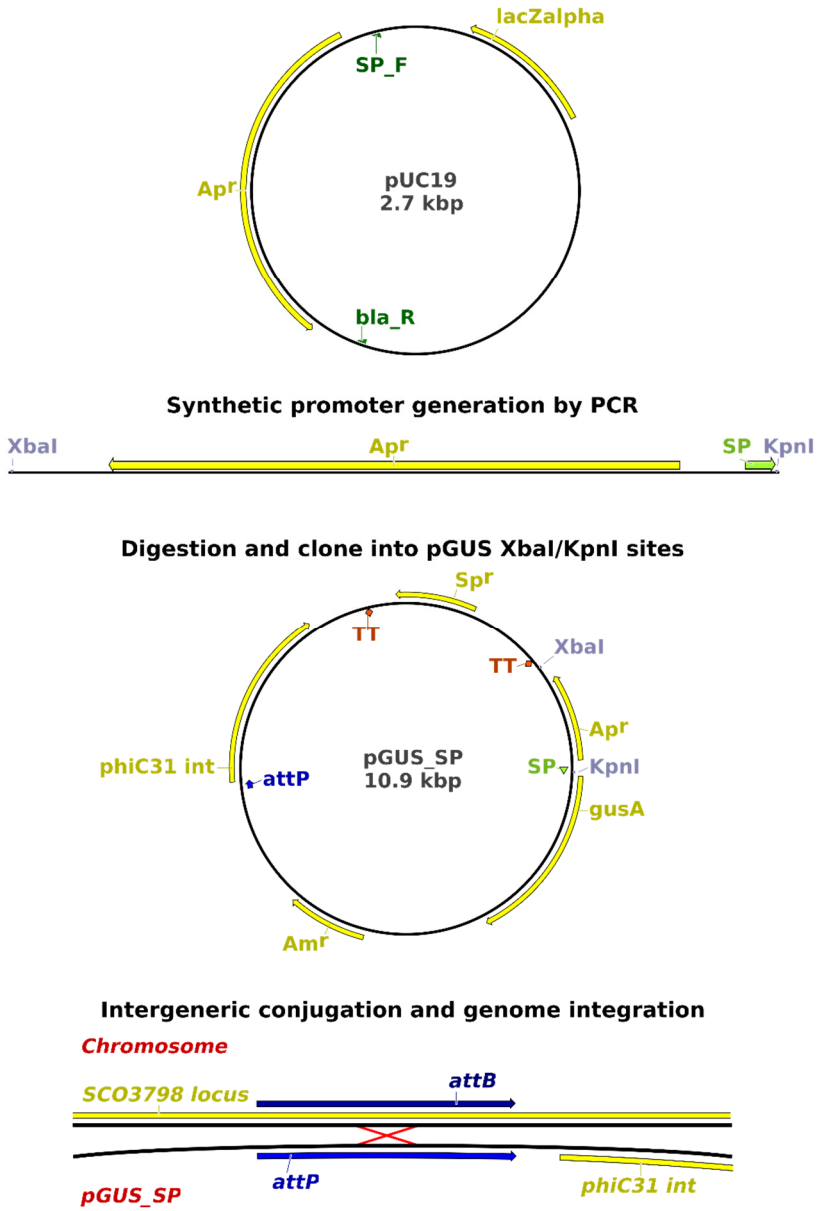


Figure 12. Strategy for construction of synthetic promoters probe vectors in *Streptomyces*. The synthetic promoters P1, P2, P3 and P21 (generically, SP) were generated by polymerase chain reaction using SP_F/bla_R primer pairs to amplify the ampicillin resistance marker of pUC19 plasmid. The PCR products were cloned into XbaI/KpnI sites of pGUS reporter vector upstream of promoterless *gusA* gene, to originate pGUS_SP plasmids. pGUS and pGUS_SP plasmids were transferred to *Streptomyces* spp. by intergeneric conjugation using non-methylating *E. coli* ET12567 [pUZ8002] as plasmid donors. Integration of pGUS_SP plasmids into *Streptomyces* chromosome by *Φ*C31 integrase-mediated *attB*/*attP* recombination was confirmed by PCR. Ap^r, ampicillin resistance marker; Am^r, apramycin resistance marker; Sp^r, spectinomycin resistance marker; TT, transcriptional terminator; SP, synthetic promoter.

Intergeneric conjugation

Foreign DNA was introduced in *Streptomyces* by intergeneric conjugation through cell-to-cell contact from a donor cell into a recipient cell. *E. coli* ET12567 [pUZ8002] carrying the pGUS_SP plasmids were used as donor cells (**figure 12**). 10 mL of LB medium supplemented with appropriate antibiotics and 20 mM $MgSO_4$ was inoculated with overnight-grown *E. coli* and incubated at 37 ºC to an OD_{600nm} of 0.9. Cells were washed twice with LB medium without antibiotics and resuspended in 600 µL of LB medium. 10^9 *Streptomyces* spp. spores were resuspended in 600 µL 2xYT medium [89] and incubated at 50 ºC for 10 minutes. Donors and recipients were incubated together at 28 ºC, for 20 min and then spread on MS plates supplemented with 10 mM $MgCl_2$ and grown overnight at 28 ºC. The plates were supplemented with apramycin and nalidixic acid, air dried and incubated at 28 ºC. The presence of exconjugates was verified after 7 days. Ten colonies for each exconjugate were streaked on new MS plates supplemented with apramycin and nalidixic acid for 7 days at 28 ºC.

Glucuronidase activity assays

For *in situ* detection of glucuronidase activity, exconjugates were streaked on MS plates supplemented with 25 µg/mL X-Gluc (5-bromo-4-chloro-3-indolyl β-D-glucuronide sodium salt [Sigma]) and incubated at 28 ºC for 24 hours.

Promoter strength was determined by measuring the glucuronidase reporter activity in total protein extracts of *Streptomyces* spp. The protocol was based in a previously reported method [100], with adaptations for continuous spectrophotometric measurements [101]. For relative promoter strength assessment, protein crude extracts were obtained from 1.5 mL samples of cultures grown in TSB for 40h and 72h. For heat-shock effect assessment, cells were grown in TSB at 28 ºC. After 40h of growth cultures were divided in two flasks to grow at 28 ºC or 42 ºC. Samples were obtained after 30 and 60 minutes of growth at 28 ºC or 42 ºC. For carbon source effect assessment, the samples were obtained from cultures grown in NMMP medium supplemented with 0.5% (w/v) glucose, glycerol or mannitol for 48h. Mycelia was washed once with one volume of 50 mM NaH_2PO_4/Na_2HPO_4 buffer pH 7 and resuspended in 0.5 mL GUS assay buffer [50 mM NaH_2PO_4/Na_2HPO4 buffer pH 7, 5 mM DTT, 10 mM Na_2EDTA, 0.1% (v/v) Triton X-100] and 1 mM phenylmethylsulfonylfluoride (PMSF). Cells were disrupted by sonication (Branson Sonifier® S-250A) and the lysate was recovered after centrifugation at 4 ºC, 16000x g. Protein content of crude extracts was determined using the Bio-Rad Protein Assay Kit (Bio-Rad), following the manufacturer's instructions and using bovine serum albumin as a standard. Crude extracts containing about 1 µg of total protein were

incubated at room temperature with 1 mM p-nitrophenyl β-D-glucuronide (PNPG) (Sigma) in 96-well plates. The conversion of PNPG into p-nitrophenol (PNP) was followed spectrophotometrically at 415 nm, for up to 20 minutes, using an iMark™ Microplate Absorbance Reader (Bio-Rad). One unit of glucuronidase (GUS unit) activity was defined as the amount of enzyme necessary to increase absorbance at 415 nm of 1 unit per minute, at room temperature. Glucuronidase specific activity was determined by calculating the PNP formation rate per minute per milligram of total protein.

Statistical analysis

All statistical analysis was performed using GraphPad Prism version 5.02.

Accession numbers

The accession numbers of sequences and studies used in this work are listed in **table 6**.

Table 6. Accession numbers of sequences and studies used in this work. The accession number GSE57268 refers to an expression profiling of WT *S. coelicolor* M145 at 24h and 72h of culture. The accession numbers GSE46507 and GSE46232 refer to two transcription maps of WT *S. coelicolor* M145 produced by differential and global RNA sequencing, respectively.

Data	Accession	Reference
S. coelicolor A3(2) chromosome sequence	NC_003888.3	[8]
S. lividans 1326 chromosome sequence	NZ_CM001889.1	[30]
S. coelicolor M145 RNA-seq	GSE57268	[94]
S. coelicolor M145 RNA-seq	GSE46507, GSE46232	[9]

Results

Screening of constitutively expressed genes in *S. coelicolor* M145

In order to find suitable templates for the designing of new promoters, constitutively expressed genes in *S. coelicolor* were screened using genome-wide transcriptomic and proteomic data. Two criteria were established: (i) genes ought to be stable and highly transcribed along time and (ii) proteomic data should support transcriptomic evidence. All *S. coelicolor* M145 genes were sorted according expression level according to time-lapse genome-wide transcriptomic data [94]. Only genes with variation in expression along time lower than 5% were considered. An ordered list of 300 genes was generated. In order to provide proteomics evidence for high expression, the 300 selected genes list were verified with most expressed proteins list in *S. coelicolor* M145 [11]**.** Twelve entries were present in both lists (**table 6**). Results were sorted using a score defined as the product of average transcripts counts and protein abundance index. Existence of TSS upstream of the selected genes was assessed using two RNA-seq datasets [9, 10]. Only 6 of the 12 selected genes presented a TSS. The three genes with highest score, SCO4761, SCO1947 and SCO0527, had TSS in their upstream regions and were selected for further characterization.

Table 7. Ordered list of stable and highly transcribed genes in *S. coelicolor* M145. EmPAI refers to *exponentially modified protein abundance index*, a measure of protein abundance, calculated elsewhere [11]. The score is the product of average transcript counts by the protein abundance index. TSS existence was assessed in the 500-bp region upstream each gene.

Locus	Average transcript counts	Protein abundance index (EmPAI)	Score	TSS	Description
SCO4761	3332	85,6	285219	Yes	co-chaperonin GroES
SCO1947	5317	27,98	148770	Yes	glyceraldehyde-3-phosphate dehydrogenase
SCO0527	3575,5	35,48	126859	Yes	cold shock protein
SCO4653	3500,5	33,5	117267	No	50S ribosomal protein L7/L12
SCO4702	3355,5	32,74	109859	No	50S ribosomal protein L3
SCO4655	3717	24,08	89505	No	DNA-directed RNA polymerase subunit beta
SCO0641	593,5	87,87	52151	Yes	tellurium resistance protein
SCO5736	790	58,93	46555	Yes	30S ribosomal protein S15
SCO3767	614,5	60,02	36882	No	hypothetical protein
SCO5776	1022	35,82	36608	No	glutamate binding protein
SCO1598	940,5	33,81	31798	No	50S ribosomal protein L20
SCO5595	953,5	30,14	28738	Yes	50S ribosomal protein L19

In silico characterization of promoter regions of the selected genes

Synthetic promoter design based in upstream region of the selected genes required a prior characterization of these regions. This characterization aimed to describe conditions or mutations known to influence gene expression and to annotate functional sequences, such as sequences recognized by sigma factors, operators, transcriptional terminators and ribosome binding sites. Following this characterization, upstream regions of the selected genes were downsize to core promoter sequences following three criteria: (i) a single TSS and the associated sequence recognized by HrdB should be included, (ii) at least, one motif associated with core promoters regions in *Streptomyces* should be present and (iii) known and predicted operators and transcriptional terminators should be avoided.

SCO4761 - groES

S. coelicolor A3(2) *groES* encodes a 10 kDa chaperonin [8]. GroES protein belongs to a multimeric complex, encompassing several subunits of GroES and GroEL proteins, that provides an isolated environment that facilities correct folding and assembly of some proteins. The sequence *of groES* in *S. coelicolor* A3(2) was first published by Duchêne *et al* [102]. *groES* TSSs are located 126 bp (TSS1) and 1 bp (TSS2) upstream of start codon [9, 102]. A putative intrinsic transcriptional terminator concerning the gene upstream *groES* (ΔG=-20.83 kcal.mol^{-1}) and a RBS (q-value=0.05) were found 269 bp and 17 bp upstream of *groES* start codon, respectively. The core promoter upstream TSS1 contains a sequence recognized by HrdB and two copies of CIRCE (controlling inverted repeat of chaperone expression), a negative *cis*-element recognized by the heat-inducible transcriptional repressor HrcA [102, 103]. Heat shock conditions were shown to induce *groES* transcription in *S. coelicolor* A3(2) and *S. lividans* TK21 [104-106]. Full exclusion of CIRCE from *dnaK* promoter in *Lactococcus lactis* was shown to abolish its heat shock-dependent transcription [107]. Furthermore, introduction of point mutations in CIRCE of *Bacillus subtilis dnaK* promoter was shown to de-repress *dnaK* expression under non-heat shock conditions [108]. Hence, in order to design a heat shock-independent synthetic promoter, CIRCE motifs should to be trimmed/removed from *groES* promoter. In *S. coelicolor* A3(2), the core promoter (upstream from TSS1) of *groES/groEL* operon contains two CIRCE: one near TSS1 (CIRCE_B), similar to *B. subtilis*, and one surrounding the -35 hexamer (CIRCE_A). CIRCE_B motif would act like in *B. subtilis*. Nevertheless, deletion of CIRCE_B would not interfere with core promoter sequence upstream TSS1. On the other hand, -35 hexamer of TSS1 was embedded into CIRCE_A motif. Thus, CIRCE_A could not be

fully excluded from the synthetic promoter template. High pairwise identity was observed between CIRCE motifs of *S. coelicolor groES* and *Bacillus subtilis dnaK* promoter (**figure 13**). Hence, together with CIRCE_B full trimming, a disruption strategy similar to IR-1 mutation [108] was applied to CIRCE_A in order to remove heat-shock regulation from this synthetic promoter sequence. Other factors were shown to up-regulate *groES* transcription, namely, the growth in chitin-amended soil [109] and pH shock [110].

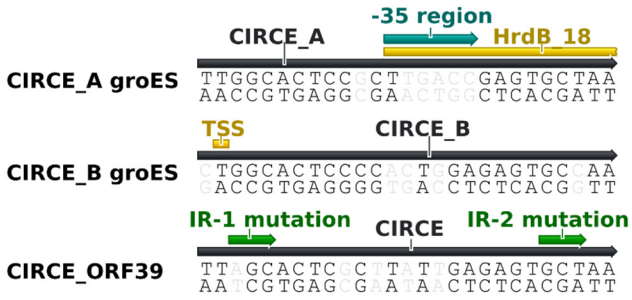


Figure 13. Alignment of CIRCE operators in promoters of *S. coelicolor* A3(2) *groES* and *Bacillus subtilis dnaK* . The IR-1/2 mutations are referred in [108].

Sequences recognized by HrdB were found upstream of TSS1 and TSS2 of *groES* [80]. The TANNNT, CANNNT and TNTNNNANNT motifs [81] were found in the -10 region of *groES* TSS1. The -35 region of *groES* TSS1 was conform to TTGAC motif [81]. No promoter core motifs compatible with TSS2 were found. Considering the annotated elements, the 43-bp region upstream to TSS1 of *S. coelicolor groES* was chosen as template for synthetic promoter and it was named P3 (**figure 14**).

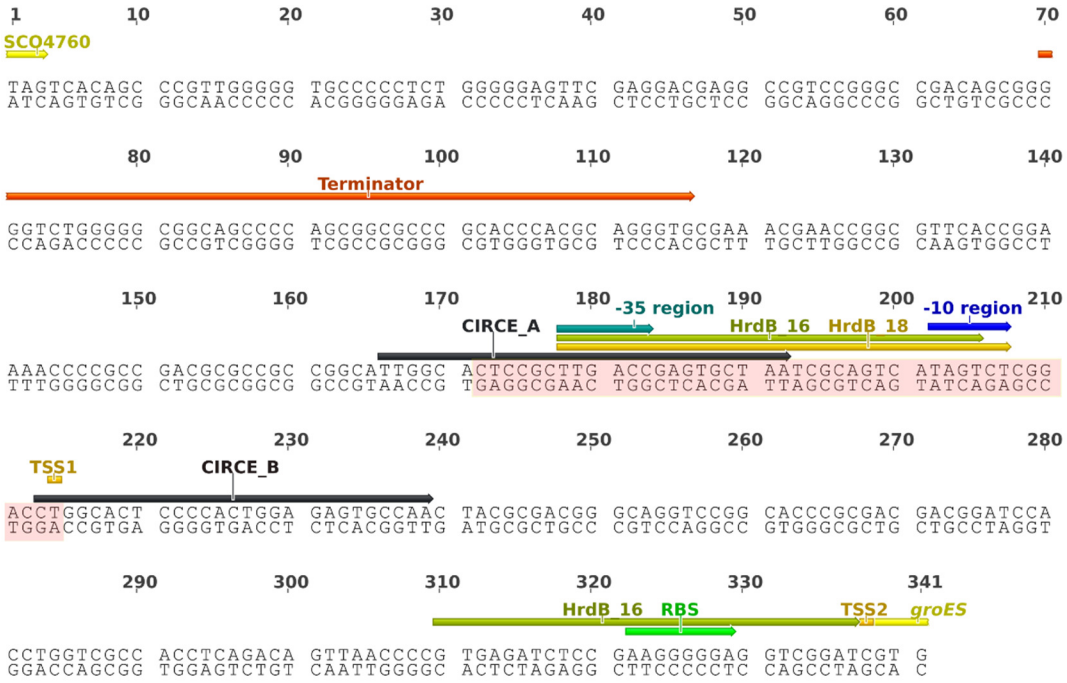


Figure 14. Annotation of the upstream region of *groES* in *S. coelicolor* A3(2). P3 promoter sequence is highlighted in pink.

SCO1947 - gap1

S. coelicolor A3(2) *gap1* encodes a glyceraldehyde-3-phosphate dehydrogenase (GAPDH) [8]. GAPDH is responsible for the reversible conversion of 1,3-bisphosphoglycerate and glyceraldehyde-3-phosphate, a central step in glycolysis and gluconeogenesis. GAPDH has been classified as constitutively expressed [111]. Despite that, Δ*SCO2179* mutant (lacking a leucyl aminopeptidase) was shown to down-regulate GAPDH [112], while glucose (in a glucose kinase-independent way) [113] and pH shock [114] were shown to up-regulate GAPDH expression. The TSSs of *gap1* are located 200 bp (TSS1) and 67 bp (TSS2) upstream of *gap1* start codon [9]. A putative intrinsic transcriptional terminator concerning the gene upstream *gap1* (ΔG=-24.49 kcal.mol^{-1}) and a RBS (q-value=0.06) were found 162 bp and 14 bp upstream of *gap1*, respectively. TSS1 was thought to be less important for *gap1* translation since it is located upstream the predicted transcriptional terminator.

The region upstream *gap1* has been associated with several operators and sequences recognized by sigma factors. A sequence recognized by sigma factor B (SigB) was located 119 bp upstream of *gap1* [115]. However, no TSS compatible with SigB-directed transcription of *gap1* has been described. AfsQ1 operator was located 11 bp upstream of *gap1*, in the nontemplate strand [116]. However, *gap1* transcription was not affected by inactivation of the AfsQ1/Q2 regulator [116]. No ARG box [117] was identified in *gap1* upstream region of *S. coelicolor* A3(2) despite of *gap1* down-regulation in *S. coelicolor* Δ*argR* mutant [118]. Three operators of the glycerol-inducible repressor [119] (GylR, **figure 15**) were found (p<0.05) in upstream region of *gap1*, one of them immediately upstream from TSS1.

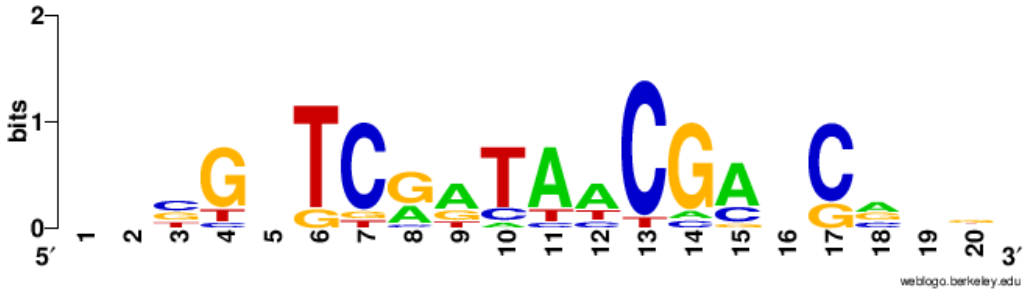


Figure 15. GylR operator in *S. coelicolor* A3(2). The GylR operator motif was reconstructed from the motifs found in the promoters of *gylR* and *glyCABX* operon [119]. The logo representation was obtained using Weblogo [97].

A 3-fold down-regulation of *gap1* transcription was observed in *S. coelicolor* Δ*bldD* mutant [120], despite no BldD operator has been described in the upstream region of *gap1*. Others factors were shown to affect *gap1* transcription, namely, differentiation stage of mycelia [121] and growth in chitin-amended soil [109].

Sequences recognized by HrdB were found upstream TSS1 and TSS2 of *gap1* [80]. The TANNNT motif [81] was found in the -10 region of *gap1* TSS2. No promoter core motifs compatible with TSS1 were found. Considering the annotated elements, the 51-bp region upstream to TSS2 of *S. coelicolor gap1* was chosen as template for synthetic promoter and it was named P2 (**figure 16**).

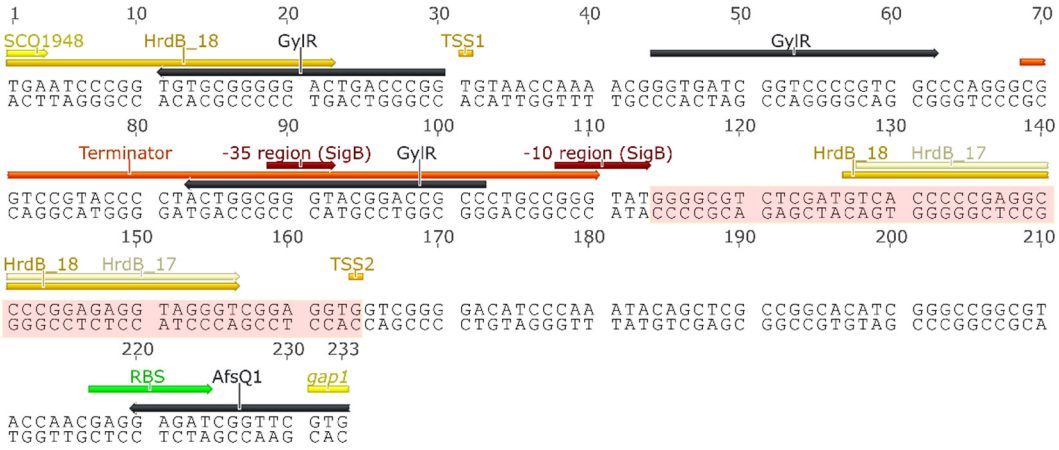


Figure 16. Annotation of the upstream region of *gap1* in *S. coelicolor* A3(2). P2 promoter sequence is highlighted in pink.

SCO0527 - scoF

S. coelicolor A3(2) *scoF* encodes for a cold shock protein, containing a RNA-binding domain that functions as RNA-chaperone [8]. The ScoF homologs in *Myxococcus xanthus*, also a soil-dwelling high-GC bacteria, are constitutively expressed at its optimal, but also at lower and higher than optimal temperatures for growth [122]. The TSS of *scoF* was located 131 bp upstream the start codon [9]. No transcriptional terminator nor RBS were predicted in the upstream region of *scoF.* Constant transcriptional levels of *scoF* were described under different pH conditions and along time of culture, and, for that reason, it has being used as housekeeping gene in RT-qPCR analysis [110, 114]. *S. coelicolor* Δ*dasR* mutant grown in chitin-amended soil was shown to down-regulate *scoF* [109]. A DasR-responsive element (dre) has been predicted 301 bp upstream of *scoF* start codon [123].

Sequences recognized by HrdB were found upstream *scoF* TSS [80]. The GTNNANNNT and TANNNT motifs [81] were found in the -10 region of *scoF* TSS. Considering the annotated elements, the 52-bp region upstream TSS of *S. coelicolor scoF* was chosen as template for synthetic promoter and it was named P1 (**figure 17**).

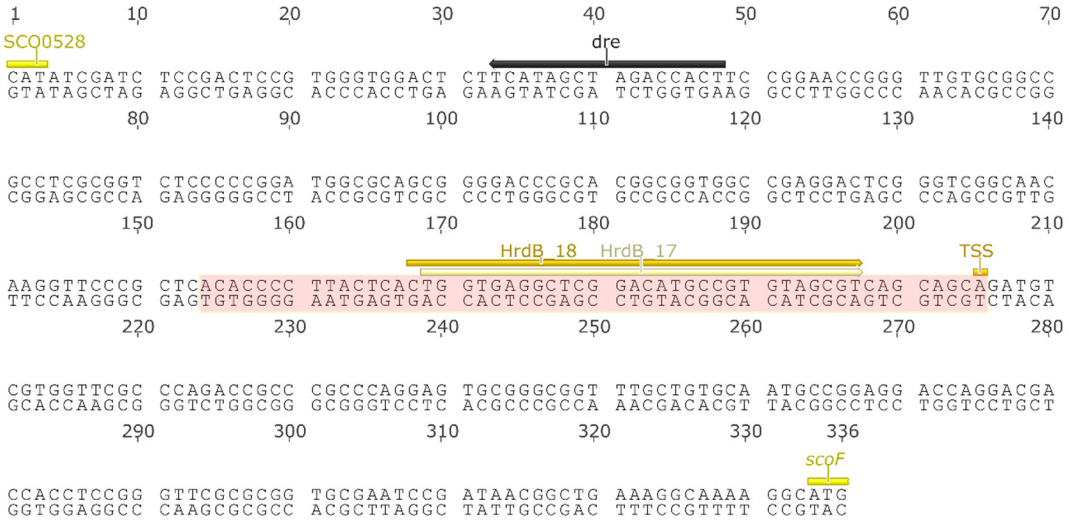


Figure 17. Annotation of the upstream region of *scoF* in *S. coelicolor* A3(2). P1 promoter sequence is highlighted in pink.

In vivo implementation of designed synthetic promoters

Three synthetic promoters - P1, P2 and P3 - based in *S. coelicolor* A3(2) *scoF*, *gap1* and *groES* native promoter regions, respectively, were designed according the available information. P21 promoter [87] was selected as reference for further *in vivo* characterizations. The synthetic promoters P1, P2, P3 and control P21 were generated by PCR and cloned upstream *gusA* reporter gene in pGUS probe vector. Promoterless *gusA* of pGUS vector was used as negative control. All plasmids were introduced in *S. coelicolor* M145 *and S. lividans* 1326 by intergeneric conjugation. Exconjugates were selected based on apramycin resistance marker of pGUS and cultured to axenic conditions. *Streptomyces* exconjugates were first tested *in situ* for β-glucuronidase (GUS) activity. Most exconjugates containing a synthetic promoter upstream *gusA* produced a blue halo of 5,5'-dibromo-4,4'-dichloro-indigo (**figure 18**), as result of β-glucuronidase activity.

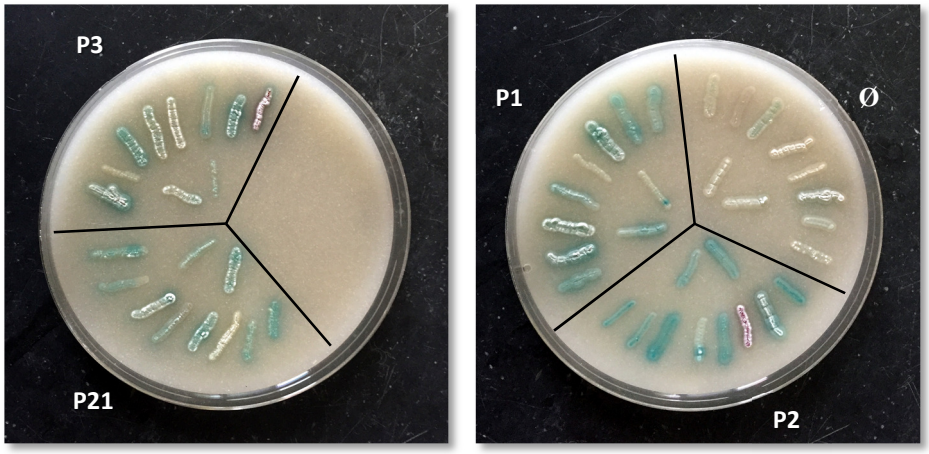


Figure 18. *In situ* detection of GUS activity in *Streptomyces* exconjugates. *S. coelicolor* pGUS or pGUS_SP (P1, P2, P3 or P21) exconjugates were streaked in MS medium supplemented with apramycin and 5-bromo-4-chloro-3-indolyl β-D-glucuronide (X-Gluc). The blue halos are indicative of β-glucuronidase activity.

Correct chromosomal integration of pGUS-derived plasmids was verified by PCR using genomic DNA of *S. coelicolor* and *S. lividans* exconjugates as template. The *Φ*C31 *attB* locus is located in a conserved pirin-like-encoding gene in *S. coelicolor* A3(2) and *S. lividans* 1326 (loci SCO3798 and SLI_RS19295, respectively). Hence, for both organisms, a primer-pair was used to confirm correct integration of plasmid into *Φ*C31 *attB* locus (C31_int_F2/R2; 2.0 kbp) and a second primer-pair to confirm the presence of synthetic promoter sequences (GUSseq_F2/pUCR; 3.2 kbp for pGUS; 4.4 kb for pGUS_SP). The expected genomic reorganization of *S. coelicolor* chromosome after pGUS_SP integration is depicted in **figure 19**.

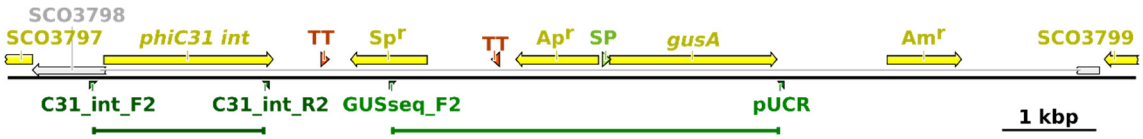


Figure 19. Expected genomic organization of *ΦC31 attB* locus of *S. coelicolor* A3(2) after pGUS_SP integration. Whole coding sequences are represented by yellow arrows. Disrupted SCO3798 gene (where *attB* was located) is represented by an interrupted silver arrow. Integration-confirming primer pairs and amplicons are depicted in green. Ap^r, ampicillin resistance marker; Am^r, apramycin resistance marker; Sp^r, spectinomycin resistance marker; TT, transcriptional terminator; SP, synthetic promoter (P1, P2, P3 or P21).

PCR amplification confirmed the genetic identity of the exconjugates (**figure 20**). Noteworthy, the appearance of a 2.5-kbp unspecific band that was shown to be unrelated to the recombination events (**figure 21**).

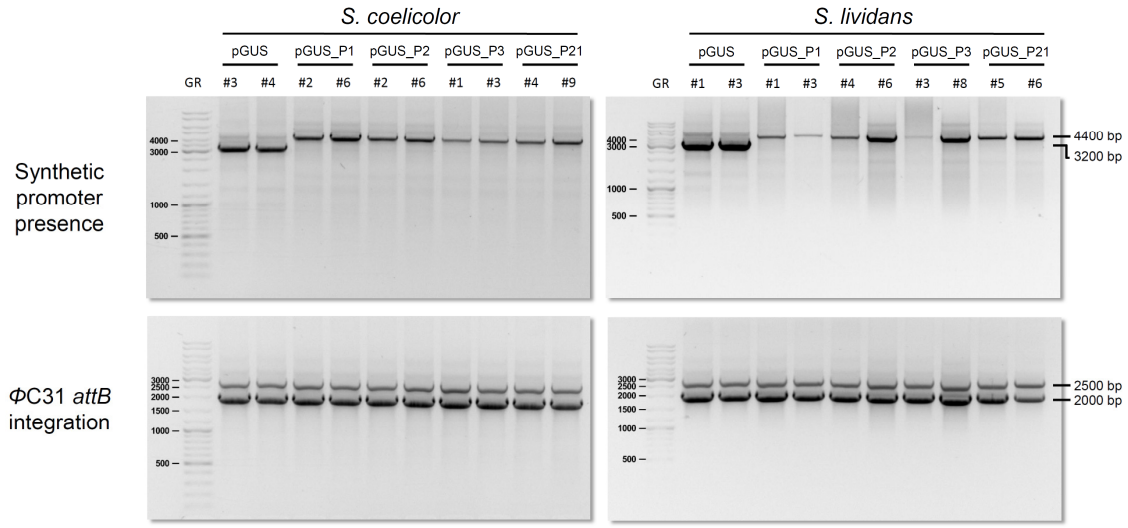


Figure 20. Gel electrophoresis of amplicons related with pGUS_SP integration into *Streptomyces* spp. genome. The amplicons were produced by PCR. Genomic DNA of the mentioned *Streptomyces* spp. exconjugates was used as template. GUSseq_F2/pUCR and C31_int_F2/C31_int_R2 primer pairs were used for synthetic promoter presence and for *ΦC31 attB* integration confirmation, respectively. Electrophoresis conditions: 1% agarose, 0.2 µg/mL ethidium bromide, 100 V, 1X TAE. GR, GeneRuler™ DNA Ladder.

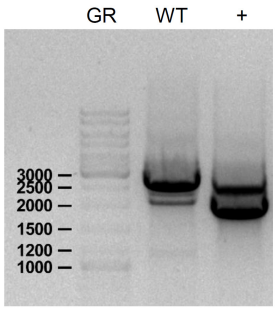


Figure 21. Unspecific band in *attB*/*attP* recombination PCR-based confirmation. The template used was genomic DNA isolated from the wild-type S. coelicolor M145 (**WT**) or from previously confirmed exconjugate of *S. coelicolor* M145 with *ΦC31 attB* integrative plasmid (pSET152, **+**). The C31_int_F2/C31_int_R2 primer pair was used for *attB* integration confirmation. Electrophoresis conditions: 1% agarose, 0.2 µg/mL ethidium bromide, 100 V, 1X TAE.

In vivo characterization of the synthetic promoters using *gusA* reporter gene

The designed synthetic promoters P1, P2, P3 and P21 were characterized in *S. coelicolor* and *S. lividans* using pGUS vector. In order to define the growth curve of exconjugates, dry weight was assessed during culture (**figure 22**). Exconjugates containing a synthetic promoter upstream *gusA* were found to replicate the growth curve of control cultures (pGUS vector; Ø). The 40h and 72h culture periods were established as early and late stationary growth phase in *S. coelicolor* and *S. lividans* exconjugates.

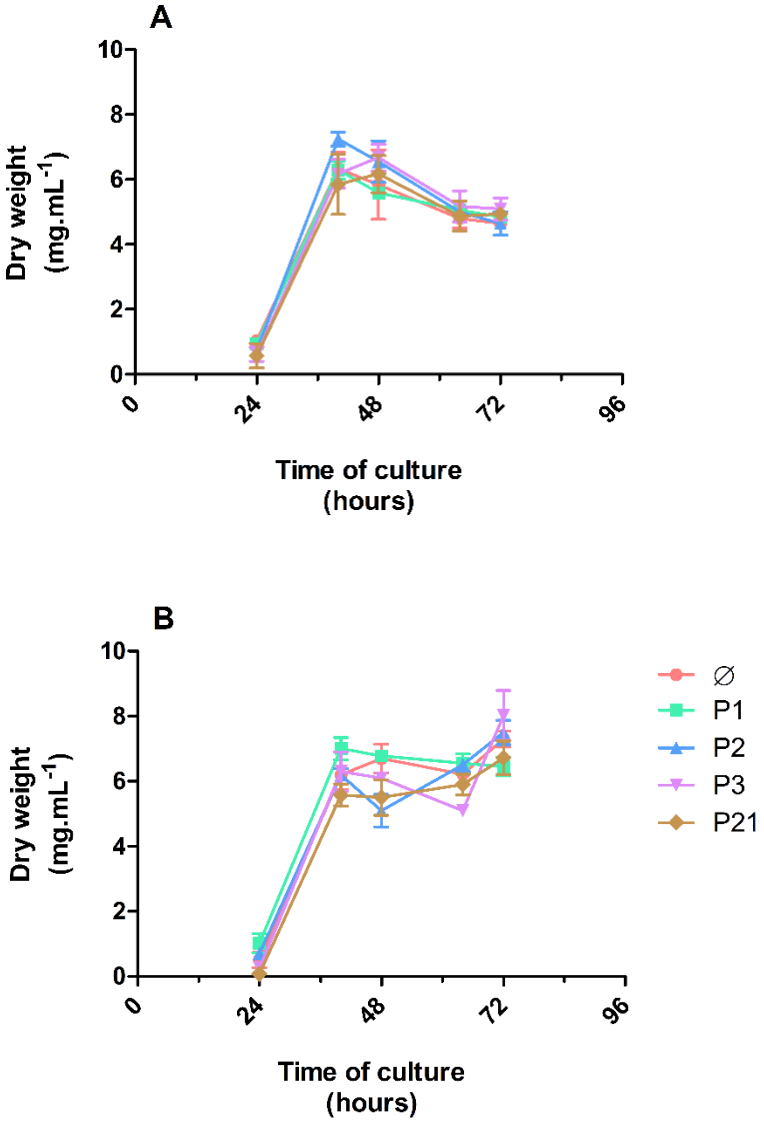


Figure 22. Growth curve of *Streptomyces* spp. exconjugates. *S. coelicolor* M145 (**A**) and *S. lividans* 1326 (**B**) exconjugates of P1, P2, P3 and P21 were grown in TSB medium, at 28 ºC. Growth was assessed by dry weight. Exconjugates of empty pGUS vector are represented by Ø. Data represent the mean and standard deviation of two technical replicates of up to two biological replicates (N=2-4).

Relative promoter strength

In order to characterize the synthetic promoters *in vivo,* β-glucuronidase (GUS) activity was assessed in cell extracts of *S. coelicolor* and *S. lividans* exconjugates in early and late stationary growth phases (**figure 23**). The GUS activity of empty vector exconjugates was negligible. P1 and P2 promoters presented similar relative strengths, ranging between 7-13% and 12-15% of P21 activity, respectively (P1: 13% at 40h, 10% at 72h for *S. coelicolor*; 8% at 40h, 7% at 72h for *S. lividans*; P2: 15% at 40h, 12% at 72h for *S. coelicolor*; 12% at 40h, 13% at 72h for *S. lividans*). P3 promoter was the strongest promoter designed in this study, ranging between 21-36% of P21 activity (P3: 21% at 40h, 22% at 72h for *S. coelicolor*; 24% at 40h, 36% at 72h for *S. lividans*). Considering the designed synthetic promoters, only P3, at 72h of culture, in *S. coelicolor* reached a statistically significant promoter activity. Interestingly, the relative activity of P1, P2 and P3 comparing to P21 was similar in *S. coelicolor* and *S. lividans*. GUS activity detected at 72h was approximately 2-fold higher than at 40h time-point in both strain, for all promoters.

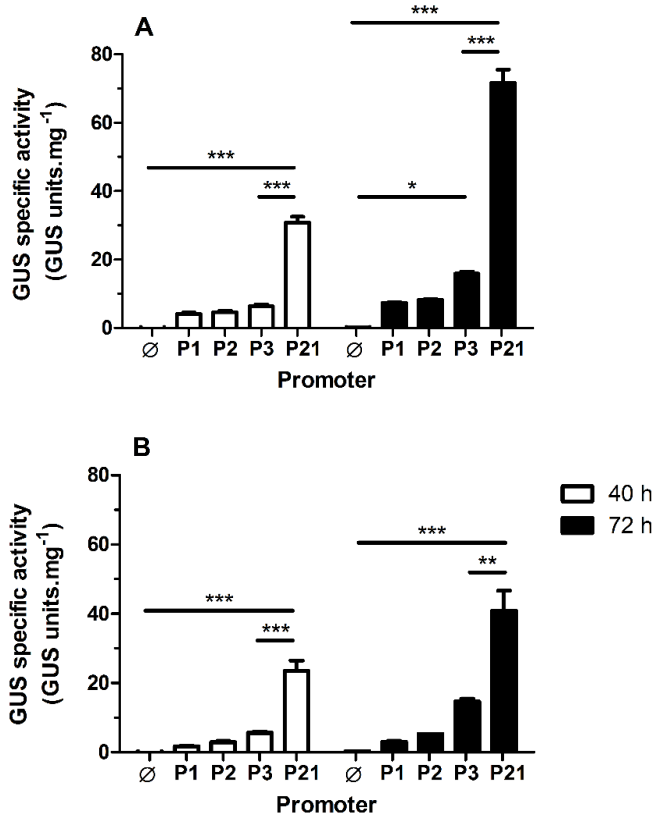


Figure 23. Glucuronidase specific activity of P1, P2, P3 and P21 promoters in *Streptomyces* spp. *S. coelicolor* M145 (**A**) and *S. lividans* 1326 (**B**) exconjugates were grown in TSB medium for 40h and 72h, at 28 ºC. GUS specific activity was assessed as the conversion rate of p-nitrophenyl β-D-glucuronide into p-nitrophenol per milligram of total protein of cell extracts. Exconjugates of empty pGUS vector are represented by Ø. Data represent the mean and standard error of the mean of two technical replicates of up to two biological replicates (N=2-4). Statistical analysis was performed using one-way ANOVA with Bonferroni's post test. *** p<0.001, ** p<0.01, * p<0.05

Heat-shock effect on P3 promoter

During P3 design, the CIRCE operators of *groES* promoter were excluded or partly trimmed to avoid heat-shock regulation. In order to assess P3 activity under heat-shock conditions, GUS activity was assessed in cell extracts of *S. coelicolor* exconjugates maintained at 28 ºC or exposed to 42 ºC for 30 and 60 minutes (**figure 24**)**.** Both P3 and P21 activities were significantly higher in cultures exposed to 42 ºC for 60 minutes relatively to basal conditions and to cultures maintained at 28 ºC during the same period of time. Hence, both promoters seem to be up-regulated by heat-shock.

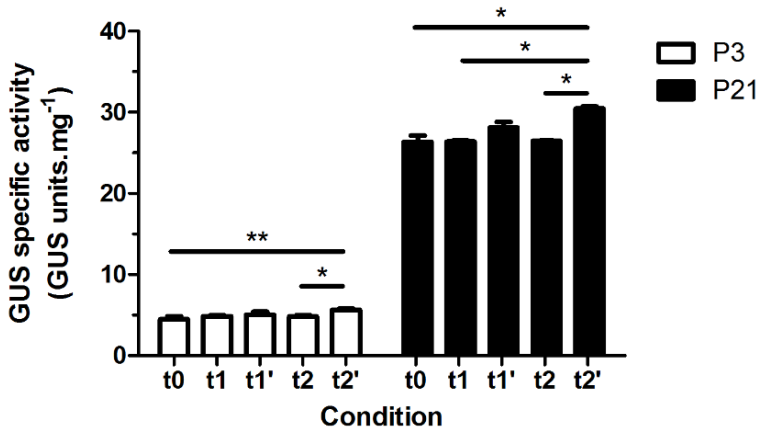


Figure 24. Glucuronidase specific activity of P3 and P21 promoters at different culture temperatures. *S. coelicolor* exconjugates were initially grown in TSB medium for 40h, at 28 ºC. GUS activity in cell extracts was assessed at 40h (**t0**), after 30 min at 28 ºC (**t1**) or 42ºC (**t1'**) and after 60 min at 28 ºC (**t2**) or 42 ºC (**t2'**). Data represent the mean and standard error of the mean of two technical replicates of up to two biological replicates (N=2-4). Statistical analysis was performed using one-way ANOVA with Bonferroni's post test. ** $p<0.01$, * $p<0.05$

Carbon source effect on P2 promoter

P2 was designed using *gap1* core promoter as template. Despite being regarded as constitutively expressed, many factors were found to influence *gap1* transcription, including carbon sources. In order to assess the influence of glucose, glycerol and mannitol on P2 activity, *S. coelicolor* exconjugates were grown in minimal medium supplemented with each carbon source separately. The growth curve of exconjugates was assessed for each carbon source by measuring dry weight during culture (**figure 25**).

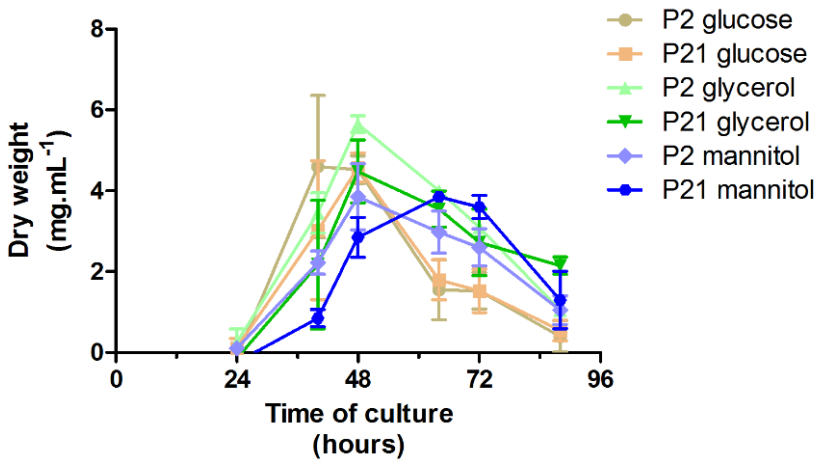


Figure 25. Growth curves of P2 and P21 *Streptomyces* exconjugates in minimal medium with defined carbon sources. *S. coelicolor* exconjugates were grown in minimal medium (NMMP) with 0.5% (w/v) glucose, glycerol or mannitol at 28 ºC. Growth was assessed by dry weight. Data represent the mean and standard deviation of two technical replicates of up to two biological replicates (N=2-4).

Lower dry weight and delayed entry in stationary grown phase were observed in cultures grown in minimal medium relatively to complex medium (TSB). Furthermore, biomass was found to suddenly decrease after maximal dry weight was reached, which generally happened at 48h of culture. GUS activity in the cell extracts was assessed at 48h of culture (**figure 26**). The GUS activities of P2 and P21 in minimal medium were about 35% lower than in TSB. No significant differences were observed in P21 activities regardless of carbon source used. P2 activity was found to be statistically different between mannitol- and glucose-supplemented minimal medium.

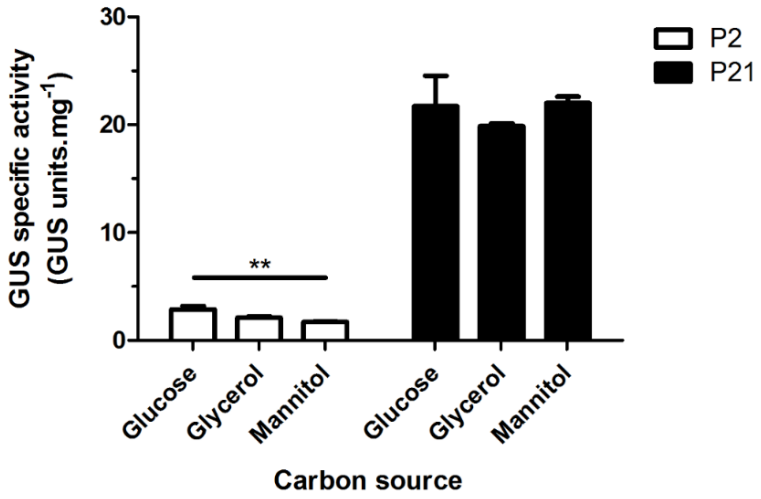


Figure 26. Glucuronidase specific activity of P2 and P21 promoters with defined carbon sources. *S. coelicolor* exconjugates were grown in minimal medium (NMMP) with 0.5% (w/v) glucose, glycerol or mannitol for 48h, at 28 ºC. GUS specific activity was assessed as conversion rate of p-nitrophenyl β-D-glucuronide into p-nitrophenol per milligram of total protein of cell extracts. Data represent the mean and standard error of the mean of two technical replicates of up to two biological replicates (N=2-4). Statistical analysis was performed using one-way ANOVA with Bonferroni's post test. ** $p<0.01$

Discussion

The development of synthetic parts is a major requirement for successful engineering in *Streptomyces*. In this work, genome-wide transcriptomic and proteomics studies of *S. coelicolor* M145 were combined with bioinformatics tools in order to select suitable templates for the design of minimal constitutive promoters.

Transcription regulation in *Streptomyces* has been shown to be far more complex than in model organism, *E. coli*, especially considering the high number of sigma factors and transcriptional regulators. Hence, *in silico* characterization and development of promoter for *Streptomyces* requires alternative approaches. In this work, core promoters sequences were established according to transcription start site (TSS) positions described in RNA-seq studies [9, 10] and to the presence of sequences recognized by HrdB [80] and short motifs predominant in *Streptomyces* promoters [81]. Using this approach, we have designed 3 synthetic promoters based in the sequences of genes screened as stable and highly expressed in *S. coelicolor* A3(2).

A reporter vector encoding β-glucuronidase protein (GUS) was used to test the designed synthetic promoters in *Streptomyces*. A high-throughput procedure was set up to determine GUS activity in cell extracts, using continuous spectrophotometric measurements adapted to 96-well plate layouts. This procedure retains several advantages: (i) rapid preparation and short execution time, (ii) micro-scale consumption of reagents, (iii) robust linear activity records and (iv) simultaneous readings of different conditions.

GUS specific activity associated to all promoters tested had similar fold increase between 40h and 72h. Constitutive activity of the tested promoters might explain the accumulation of GUS protein during time of culture. In fact, *S. coelicolor* extracts would retain 50% of GUS specific activity approximately 7 hours after transcription de-induction [100]. Therefore, the synthetic promoters generated in this work were able to direct transcription in *S. coelicolor* and *S. lividans* as stably as a known constitutive promoter, in complex media and standard conditions of temperature and aeration.

The new synthetic promoters displayed significantly lower activity than P21. P1, based in *S. coelicolor scoF* core promoter, reached 7-13% of P21 activity (about 17% of *ermE*p*). Recently, the use of full upstream regions of *scoF* in *S. albus* (XN_3584) was shown to direct 4-fold higher activity than *ermE*p* [124]. However, P1 and the referred sequence in *S. albus* present only 54.2% pairwise identity. The absence of *dre* operator

from P1 might have discarded DasR activator activity observed in *scoF* and contributed to its lower promoter activity.

P2 promoter, based in *S. coelicolor gap1* core promoter, achieved 12-15% of P21 activity (about 17% of *ermE*p*). Recently, the use of full upstream regions of *gap1* of *S. griseus* and *S. albus* J1074 was shown to direct up to 5-fold higher activity than *ermE*p* [88, 124]. The core promoter sequence used as P2 template was shown 83.7% pairwise identity in the regions upstream *gap1* of *S. coelicolor, S. griseus* and *S. albus*, including -10 and -35 regions and the spacer length. Neither the sequence recognized by SigB nor the region upstream of *gap1* TSS1 in *S. coelicolor* present such pairwise identity with *S. griseus* or *S. albus* sequences (40.8% and 20%, respectively). Interestingly, high pairwise identity was found in the RBS of *gap1* in *S. coelicolor*, *S. griseus* and *S. albus*. This suggest that RBS of *gap1* might play an important role on native and reporter protein levels [88, 124]. Therefore, the lower activity of P2 relatively to *gap1* promoter might be related with dissimilar translation rates rather than different promoter activities. The activity of P2 was shown statistically different in mannitol- versus glucose-containing medium. Mannitol, as non-repressing carbon source, was not expected to down regulate P2 activity. In fact, a 3-fold increase of GAPDH expression was observed in *S. coelicolor* M145 grown in glucose and mannitol versus grown in mannitol only [113]. Hence, the correct interpretation of the data might be that glucose is up-regulating P2. Glycerol was shown not to influence P2 activity. Thus, either GylR was not a *gap1* regulator in *S. coelicolor* or the exclusion of the detected GlyR operators was sufficient to remove glycerol dependence. Intriguingly, at stationary phase, P2 and P21 shown lower activity in NMMP-glucose than in TSB. This observation could indicate an additional layer of regulation of these promoters, concisely related to stringent response [125]. Further stringent response-inductive experiments would be required to test this hypothesis.

P3, based in *S. coelicolor groES* core promoter, achieved 21-36% of P21 activity (about 33% of *ermE*p*). Having the most HrdB-compliant sequence and core promoter motifs both in -10 and -35 regions, P3 was also the most active of the synthetic promoters design in this work. High promoter activity (4-fold higher than *ermE*p*) was recently attributed to full region upstream *groES* in *S. albus* J1074 [124]. Broad pairwise identity was found between the upstream regions of *groES* in *S. coelicolor* and *S. albus*, including the core promoter region upstream TSS1 and the CIRCE motifs. Interestingly, the non-trimmed region upstream of *groES* in *S. albus* retained high promoter activity in non-heat-shock conditions. The alternative transcription of *groES* starting at TSS2, which is absent from P3, might explain the disparities in promoter activity observed.

The stability of P3 was determined in heat-shock conditions. Regardless of promoter tested, P3 or P21, a generalized increase of promoter activity was observed in heat-shock conditions. This behaviour could be partly explained by acceleration of metabolic rate due to temperature shift [126]. On the other hand, the heat-shock induction of P3 and P21 suggest not only P3 might retained the heat-shock regulation of *groES* promoter but also that P21 could be inducible by heat-shock.

In silico and *in vivo* characterization of promoter regions should not miss functional elements and conditions known to influence their transcriptional activity. Recently, large sets of full upstream regions of constitutively expressed genes in *Streptomyces* were described as constitutive promoters [88, 124]. However, not only the presence of multiple sources of genetic expression conditioning, such as transcriptional terminators, multiple TSS and native RBS, but also insufficient characterization prevent those full upstream regions from attaining the desired reliability for synthetic biology approaches. In this work, the designed synthetic promoters were investigated regarding the regulation of their template promoters, guaranteeing the predictability required in new parts for genetic engineering.

Conclusions

Publically available genome-wide transcriptomic and proteomics data can be used to select stable and highly expressed genes in *S. coelicolor*. *In silico* characterization of promoter regions of stable and highly expressed genes can be applied to design functional synthetic promoter for *Streptomyces*. A comprehensive *in vivo* characterization of synthetic promoters should include conditions known to affect the native promoter region.

Future perspectives

Further confirmation of transcriptional activity of the designed synthetic promoters would be performed by RT-qPCR. Synthetic ribosome binding sites could be attached to the designed promoters in order to study genetic expression modulation at translational level. The promoters designed in this study could be tested in a plug-and-play scaffold for gene cluster refactoring.

References

1. Goodfellow, M., et al., *Bergey's Manual of Systematic Bacteriology: Volume 5: The Actinobacteria*. 2012: Springer New York.
2. Euzeby, J.P., *List of Bacterial Names with Standing in Nomenclature: a folder available on the Internet.* Int J Syst Bacteriol, 1997. **47**(2): p. 590-2.
3. Lucas, X., et al., *StreptomeDB: a resource for natural compounds isolated from Streptomyces species.* Nucleic Acids Res, 2013. **41**(Database issue): p. D1130-6.
4. Chater, K.F., *Streptomyces inside-out: a new perspective on the bacteria that provide us with antibiotics.* Philosophical Transactions of the Royal Society of London B: Biological Sciences, 2006. **361**(1469): p. 761-768.
5. Hopwood, D.A., *Forty years of genetics with Streptomyces: from in vivo through in vitro to in silico.* Microbiology, 1999. **145 (Pt 9)**: p. 2183-202.
6. Bentley, S.D., et al., *SCP1, a 356,023 bp linear plasmid adapted to the ecology and developmental biology of its host, Streptomyces coelicolor A3(2).* Mol Microbiol, 2004. **51**(6): p. 1615-28.
7. Haug, I., et al., *Streptomyces coelicolor A3(2) plasmid SCP2*: deductions from the complete sequence.* Microbiology, 2003. **149**(Pt 2): p. 505-13.
8. Bentley, S.D., et al., *Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2).* Nature, 2002. **417**(6885): p. 141-7.
9. Romero, D.A., et al., *A comparison of key aspects of gene regulation in Streptomyces coelicolor and Escherichia coli using nucleotide-resolution transcription maps produced in parallel by global and differential RNA sequencing.* Mol Microbiol, 2014.
10. Vockenhuber, M.P., et al., *Deep sequencing-based identification of small non-coding RNAs in Streptomyces coelicolor.* RNA Biol, 2011. **8**(3): p. 468-77.
11. Thomas, L., et al., *Metabolic switches and adaptations deduced from the proteomes of Streptomyces coelicolor wild type and phoP mutant grown in batch culture.* Mol Cell Proteomics, 2012. **11**(2): p. M111 013797.
12. Berdy, J., *Bioactive microbial metabolites.* J Antibiot (Tokyo), 2005. **58**(1): p. 1-26.
13. Tarkka, M. and R. Hampp, *Secondary metabolites of soil streptomycetes in biotic interactions*, in *Secondary metabolites in soil ecology*. 2008, Springer. p. 107-126.

14. Nett, M., H. Ikeda, and B.S. Moore, *Genomic basis for natural product biosynthetic diversity in the actinomycetes.* Nat Prod Rep, 2009. **26**(11): p. 1362-84.
15. Harrison, J. and D.J. Studholme, *Recently published Streptomyces genome sequences.* Microb Biotechnol, 2014. **7**(5): p. 373-80.
16. Medema, M.H. and M.A. Fischbach, *Computational approaches to natural product discovery.* Nat Chem Biol, 2015. **11**(9): p. 639-48.
17. Weber, T., et al., *antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters.* Nucleic Acids Res, 2015. **43**(W1): p. W237-43.
18. Baltz, R.H., *Renaissance in antibacterial discovery from actinomycetes.* Curr Opin Pharmacol, 2008. **8**(5): p. 557-63.
19. Hwang, K.S., et al., *Systems biology and biotechnology of Streptomyces species for the production of secondary metabolites.* Biotechnol Adv, 2014. **32**(2): p. 255-68.
20. Waksman, S.A., *Streptomycin: background, isolation, properties, and utilization.* Science, 1953. **118**(3062): p. 259-66.
21. Oura, M., T.H. Sternberg, and E.T. Wright, *A new antifungal antibiotic, amphotericin B.* Antibiot Annu, 1955. **3**: p. 566-73.
22. Burg, R.W., et al., *Avermectins, new family of potent anthelmintic agents: producing organism and fermentation.* Antimicrob Agents Chemother, 1979. **15**(3): p. 361-7.
23. Goto, T., et al., *Discovery of FK-506, a novel immunosuppressant isolated from Streptomyces tsukubaensis.* Transplant Proc, 1987. **19**(5 Suppl 6): p. 4-8.
24. Murakami, T., et al., *The bialaphos biosynthetic genes ofStreptomyces hygroscopicus: Molecular cloning and characterization of the gene cluster.* Molecular and General Genetics MGG, 1986. **205**(1): p. 42-53.
25. Lautru, S., et al., *Discovery of a new peptide natural product by Streptomyces coelicolor genome mining.* Nat Chem Biol, 2005. **1**(5): p. 265-9.
26. Gerber, N.N. and H.A. Lechevalier, *Geosmin, an earthly-smelling substance isolated from actinomycetes.* Appl Microbiol, 1965. **13**(6): p. 935-8.
27. Serrano, L., *Synthetic biology: promises and challenges.* Molecular Systems Biology, 2007. **3**(1): p. 158.
28. Medema, M.H., et al., *Exploiting plug-and-play synthetic biology for drug discovery and production in microorganisms.* Nat Rev Microbiol, 2011. **9**(2): p. 131-7.

29. Medema, M.H., R. Breitling, and E. Takano, *Synthetic biology in Streptomyces bacteria.* Methods Enzymol, 2011. **497**: p. 485-502.
30. Cruz-Morales, P., et al., *The genome sequence of Streptomyces lividans 66 reveals a novel tRNA-dependent peptide biosynthetic system within a metal-related genomic island.* Genome Biol Evol, 2013. **5**(6): p. 1165-75.
31. Zaburannyi, N., et al., *Insights into naturally minimised Streptomyces albus J1074 genome.* BMC Genomics, 2014. **15**: p. 97.
32. Feng, Z., et al., *Engineered production of iso-migrastatin in heterologous Streptomyces hosts.* Bioorg Med Chem, 2009. **17**(6): p. 2147-53.
33. Gomez-Escribano, J.P. and M.J. Bibb, *Engineering Streptomyces coelicolor for heterologous expression of secondary metabolite gene clusters.* Microb Biotechnol, 2011. **4**(2): p. 207-15.
34. Komatsu, M., et al., *Genome-minimized Streptomyces host for the heterologous expression of secondary metabolism.* Proc Natl Acad Sci U S A, 2010. **107**(6): p. 2646-51.
35. Beites, T. and M.V. Mendes, *Chassis optimization as a cornerstone for the application of synthetic biology based strategies in microbial secondary metabolism.* Frontiers in Microbiology, 2015. **6**.
36. Mazodier, P., R. Petter, and C. Thompson, *Intergeneric conjugation between Escherichia coli and Streptomyces species.* J Bacteriol, 1989. **171**(6): p. 3583-5.
37. Flett, F., V. Mersinias, and C.P. Smith, *High efficiency intergeneric conjugal transfer of plasmid DNA from Escherichia coli to methyl DNA-restricting streptomycetes.* FEMS Microbiol Lett, 1997. **155**(2): p. 223-9.
38. Gust, B., et al., *PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin.* Proc Natl Acad Sci U S A, 2003. **100**(4): p. 1541-6.
39. Khodakaramian, G., et al., *Expression of Cre recombinase during transient phage infection permits efficient marker removal in Streptomyces.* Nucleic Acids Res, 2006. **34**(3): p. e20.
40. Horvath, P. and R. Barrangou, *CRISPR/Cas, the immune system of bacteria and archaea.* Science, 2010. **327**(5962): p. 167-70.
41. Huang, H., et al., *One-step high-efficiency CRISPR/Cas9-mediated genome editing in Streptomyces.* Acta Biochim Biophys Sin (Shanghai), 2015. **47**(4): p. 231-43.
42. Cobb, R.E., Y. Wang, and H. Zhao, *High-efficiency multiplex genome editing of Streptomyces species using an engineered CRISPR/Cas system.* ACS Synth Biol, 2015. **4**(6): p. 723-8.

43. Lydiate, D.J., F. Malpartida, and D.A. Hopwood, *The Streptomyces plasmid SCP2*: its functional analysis and development into useful cloning vectors.* Gene, 1985. **35**(3): p. 223-35.
44. Bierman, M., et al., *Plasmid cloning vectors for the conjugal transfer of DNA from Escherichia coli to Streptomyces spp.* Gene, 1992. **116**(1): p. 43-9.
45. Gregory, M.A., R. Till, and M.C. Smith, *Integration site for Streptomyces phage phiBT1 and development of site-specific integrating vectors.* J Bacteriol, 2003. **185**(17): p. 5320-3.
46. Hong, H.J., et al., *The role of the novel Fem protein VanK in vancomycin resistance in Streptomyces coelicolor.* J Biol Chem, 2005. **280**(13): p. 13055-61.
47. Myronovskyi, M., et al., *Beta-glucuronidase as a sensitive and versatile reporter in actinomycetes.* Appl Environ Microbiol, 2011. **77**(15): p. 5370-83.
48. Ingram, C., et al., *xylE functions as an efficient reporter gene in Streptomyces spp.: use for the study of galP1, a catabolite-controlled promoter.* J Bacteriol, 1989. **171**(12): p. 6617-24.
49. Liu, H., et al., *Rapid cloning and heterologous expression of the meridamycin biosynthetic gene cluster using a versatile Escherichia coli-streptomyces artificial chromosome vector, pSBAC.* J Nat Prod, 2009. **72**(3): p. 389-95.
50. Andrianantoandro, E., et al., *Synthetic biology: new engineering rules for an emerging discipline.* Mol Syst Biol, 2006. **2**: p. 2006 0028.
51. Bibb, M.J., G.R. Janssen, and J.M. Ward, *Cloning and analysis of the promoter region of the erythromycin resistance gene (ermE) of Streptomyces erythraeus.* Gene, 1985. **38**(1-3): p. 215-26.
52. Bibb, M.J., et al., *The mRNA for the 23S rRNA methylase encoded by the ermE gene of Saccharopolyspora erythraea is translated in the absence of a conventional ribosome-binding site.* Mol Microbiol, 1994. **14**(3): p. 533-45.
53. Hindle, Z. and C.P. Smith, *Substrate induction and catabolite repression of the Streptomyces coelicolor glycerol operon are mediated through the GylR protein.* Mol Microbiol, 1994. **12**(5): p. 737-45.
54. Herai, S., et al., *Hyper-inducible expression system for streptomycetes.* Proc Natl Acad Sci U S A, 2004. **101**(39): p. 14031-5.
55. Rodriguez-Garcia, A., et al., *Natural and synthetic tetracycline-inducible promoters for use in the antibiotic-producing bacteria Streptomyces.* Nucleic Acids Res, 2005. **33**(9): p. e87.
56. Takano, E., et al., *A bacterial hormone (the SCB1) directly controls the expression of a pathway-specific regulatory gene in the cryptic type I polyketide*

biosynthetic gene cluster of Streptomyces coelicolor. Mol Microbiol, 2005. **56**(2): p. 465-79.

57. Horbal, L., V. Fedorenko, and A. Luzhetskyy, *Novel and tightly regulated resorcinol and cumate-inducible expression systems for Streptomyces and other actinobacteria.* Appl Microbiol Biotechnol, 2014. **98**(20): p. 8641-55.
58. Holmes, D.J., J.L. Caso, and C.J. Thompson, *Autogenous transcriptional activation of a thiostrepton-induced gene in Streptomyces lividans.* EMBO J, 1993. **12**(8): p. 3183-91.
59. Takano, E., et al., *Construction of thiostrepton-inducible, high-copy-number expression vectors for use in Streptomyces spp.* Gene, 1995. **166**(1): p. 133-7.
60. Chiu, M.L., et al., *Broad spectrum thiopeptide recognition specificity of the Streptomyces lividans TipAL protein and its role in regulating gene expression.* J Biol Chem, 1999. **274**(29): p. 20578-86.
61. Lussier, F.X., F. Denis, and F. Shareck, *Adaptation of the highly productive T7 expression system to Streptomyces lividans.* Appl Environ Microbiol, 2010. **76**(3): p. 967-70.
62. Weisblum, B. and J. Davies, *Antibiotic inhibitors of the bacterial ribosome.* Bacteriological Reviews, 1968. **32**(4 Pt 2): p. 493-528.
63. Cadenas, R.F., J.F. Martin, and J.A. Gil, *Construction and characterization of promoter-probe vectors for Corynebacteria using the kanamycin-resistance reporter gene.* Gene, 1991. **98**(1): p. 117-21.
64. Hodge, D.R., et al., *Transcriptional activation analysis by the chloramphenicol acetyl transferase (CAT) enzyme assay.* Methods Mol Biol, 1995. **37**: p. 409-21.
65. Labes, G., M. Bibb, and W. Wohlleben, *Isolation and characterization of a strong promoter element from the Streptomyces ghanaensis phage I19 using the gentamicin resistance gene (aacC1) of Tn 1696 as reporter.* Microbiology, 1997. **143 (Pt 5)**: p. 1503-12.
66. Paget, M.S., G. Hintermann, and C.P. Smith, *Construction and application of streptomycete promoter probe vectors which employ the Streptomyces glaucescens tyrosinase-encoding gene as reporter.* Gene, 1994. **146**(1): p. 105-10.
67. Craney, A., et al., *A synthetic luxCDABE gene cluster optimized for expression in high-GC bacteria.* Nucleic Acids Res, 2007. **35**(6): p. e46.
68. Cormack, B.P., R.H. Valdivia, and S. Falkow, *FACS-optimized mutants of the green fluorescent protein (GFP).* Gene, 1996. **173**(1 Spec No): p. 33-8.

69. Han, L., et al., *Time-lapsed confocal microscopy reveals temporal and spatial expression of the lysine epsilon-aminotransferase gene in Streptomyces clavuligerus.* Mol Microbiol, 1999. **34**(5): p. 878-86.
70. Knirschova, R., et al., *Utilization of a reporter system based on the blue pigment indigoidine biosynthetic gene bpsA for detection of promoter activity and deletion of genes in Streptomyces.* J Microbiol Methods, 2015. **113**: p. 1-3.
71. Li, P., et al., *An efficient blue-white screening based gene inactivation system for Streptomyces.* Appl Microbiol Biotechnol, 2015. **99**(4): p. 1923-33.
72. Orelle, C., et al., *Protein synthesis by ribosomes with tethered subunits.* Nature, 2015. **524**(7563): p. 119-24.
73. Borovok, I., et al., *Coenzyme B12 controls transcription of the Streptomyces class Ia ribonucleotide reductase nrdABS operon via a riboswitch mechanism.* J Bacteriol, 2006. **188**(7): p. 2512-20.
74. Rudolph, M.M., M.P. Vockenhuber, and B. Suess, *Synthetic riboswitches for the conditional control of gene expression in Streptomyces coelicolor.* Microbiology, 2013. **159**(Pt 7): p. 1416-22.
75. Romero-Rodriguez, A., I. Robledo-Casados, and S. Sanchez, *An overview on transcriptional regulators in Streptomyces.* Biochim Biophys Acta, 2015. **1849**(8): p. 1017-39.
76. Stanton, B.C., et al., *Genomic mining of prokaryotic repressors for orthogonal logic gates.* Nat Chem Biol, 2014. **10**(2): p. 99-105.
77. Shimada, T., et al., *The whole set of constitutive promoters recognized by RNA polymerase RpoD holoenzyme of Escherichia coli.* PLoS One, 2014. **9**(3): p. e90447.
78. Buttner, M.J., K.F. Chater, and M.J. Bibb, *Cloning, disruption, and transcriptional analysis of three RNA polymerase sigma factor genes of Streptomyces coelicolor A3(2).* J Bacteriol, 1990. **172**(6): p. 3367-78.
79. Shiina, T., K. Tanaka, and H. Takahashi, *Sequence of hrdB, an essential gene encoding sigma-like transcription factor of Streptomyces coelicolor A3(2): homology to principal sigma factors.* Gene, 1991. **107**(1): p. 145-8.
80. Strohl, W.R., *Compilation and analysis of DNA sequences associated with apparent streptomycete promoters.* Nucleic Acids Res, 1992. **20**(5): p. 961-74.
81. Bourn, W.R. and B. Babb, *Computer assisted identification and classification of streptomycete promoters.* Nucleic Acids Res, 1995. **23**(18): p. 3696-703.
82. Hawley, D.K. and W.R. McClure, *Compilation and analysis of Escherichia coli promoter DNA sequences.* Nucleic Acids Res, 1983. **11**(8): p. 2237-55.

83. Patek, M. and J. Nesvera, *Sigma factors and promoters in Corynebacterium glutamicum.* J Biotechnol, 2011. **154**(2-3): p. 101-13.
84. Wang, W., et al., *An engineered strong promoter for streptomycetes.* Appl Environ Microbiol, 2013. **79**(14): p. 4484-92.
85. Jensen, P.R. and K. Hammer, *The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters.* Appl Environ Microbiol, 1998. **64**(1): p. 82-7.
86. Seghezzi, N., et al., *The construction of a library of synthetic promoters revealed some specific features of strong Streptomyces promoters.* Appl Microbiol Biotechnol, 2011. **90**(2): p. 615-23.
87. Siegl, T., et al., *Design, construction and characterisation of a synthetic promoter library for fine-tuned gene expression in actinomycetes.* Metab Eng, 2013. **19**: p. 98-106.
88. Shao, Z., et al., *Refactoring the silent spectinabilin gene cluster using a plug-and-play scaffold.* ACS Synth Biol, 2013. **2**(11): p. 662-9.
89. Kieser, T., et al., *Practical Streptomyces Genetics*. 2000: The John Innes Foundation.
90. Rico, S., et al., *Regulation of the AbrA1/A2 two-component system in Streptomyces coelicolor and the potential of its deletion strain as a heterologous host for antibiotic production.* PLoS One, 2014. **9**(10): p. e109844.
91. Hanahan, D., *Studies on transformation of Escherichia coli with plasmids.* J Mol Biol, 1983. **166**(4): p. 557-80.
92. Gust, B., et al., *Lambda red-mediated genetic manipulation of antibiotic-producing Streptomyces.* Adv Appl Microbiol, 2004. **54**: p. 107-28.
93. Yanisch-Perron, C., J. Vieira, and J. Messing, *Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors.* Gene, 1985. **33**(1): p. 103-19.
94. Naseer, N., J.A. Shapiro, and M. Chander, *RNA-Seq analysis reveals a six-gene SoxR regulon in Streptomyces coelicolor.* PLoS One, 2014. **9**(8): p. e106181.
95. Kearse, M., et al., *Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data.* Bioinformatics, 2012. **28**(12): p. 1647-9.
96. Mitra, A., et al., *WebGeSTer DB--a transcription terminator database.* Nucleic Acids Res, 2011. **39**(Database issue): p. D129-35.
97. Crooks, G.E., et al., *WebLogo: a sequence logo generator.* Genome Res, 2004. **14**(6): p. 1188-90.

98. Grant, C.E., T.L. Bailey, and W.S. Noble, *FIMO: scanning for occurrences of a given motif.* Bioinformatics, 2011. **27**(7): p. 1017-8.
99. Hanahan, D., *Techniques for transformation of E. coli.* DNA cloning, 1985. **1**: p. 109-135.
100. Rudolph, M.M., M.P. Vockenhuber, and B. Suess, *Conditional control of gene expression by synthetic riboswitches in Streptomyces coelicolor.* Methods Enzymol, 2015. **550**: p. 283-99.
101. Aich, S., L.T. Delbaere, and R. Chen, *Continuous spectrophotometric assay for beta-glucuronidase.* Biotechniques, 2001. **30**(4): p. 846-50.
102. Duchêne, A.M., et al., *Characterization of two groEL genes in Streptomyces coelicolor A3(2).* Gene, 1994. **144**(1): p. 97-101.
103. Pascale, S. and M. Philippe, *Negative regulation of the heat shock response in Streptomyces.* Archives of Microbiology, 2001. **176**(4): p. 237-242.
104. Duchêne, A.M., C.J. Thompson, and P. Mazodier, *Transcriptional analysis of groEL genes in Streptomyces coelicolor A3(2).* Molecular & general genetics : MGG, 1994. **245**(1): p. 61-68.
105. de León, P., et al., *Streptomyces lividans groES, groEL1 and groEL2 genes.* Microbiology (Reading, England), 1997. **143 (Pt 11)**: p. 3563-3571.
106. Bucca, G., et al., *Development and application of versatile high density microarrays for genome-wide analysis of Streptomyces coelicolor: characterization of the HspR regulon.* Genome biology, 2009. **10**(1).
107. Asseldonk, M.v., et al., *Cloning, nucleotide sequence, and regulatory analysis of the Lactococcus lactis dnaJ gene.* Journal of bacteriology, 1993. **175**(6): p. 1637-1644.
108. Zuber, U. and W. Schumann, *CIRCE, a novel heat shock element involved in regulation of heat shock operon dnaK of Bacillus subtilis.* Journal of bacteriology, 1994. **176**(5): p. 1359-1363.
109. Nazari, B., et al., *Chitin-Induced Gene Expression in Secondary Metabolic Pathways of Streptomyces coelicolor A3(2) Grown in Soil.* Applied and Environmental Microbiology, 2012. **79**(2): p. 707713.
110. Yoon Jung, K., et al., *Acidic pH shock induces the expressions of a wide range of stress-response genes.* BMC genomics, 2008. **9**: p. 604.
111. Hesketh, A.R., et al., *Primary and secondary metabolism, and post-translational protein modifications, as portrayed by proteomic analysis of Streptomyces coelicolor.* Molecular microbiology, 2002. **46**(4): p. 917-932.
112. Eunjung, S., et al., *Deletion of an architectural unit, leucyl aminopeptidase (SCO2179), in Streptomyces coelicolor increases actinorhodin production and*

sporulation. Applied Microbiology and Biotechnology, 2013. **97**(15): p. 6823-6833.

113. Jacob, G., et al., *Identification of glucose kinase-dependent and -independent pathways for carbon control of primary metabolism, development and antibiotic production in Streptomyces coelicolor by quantitative proteomics.* Molecular microbiology, 2012. **86**(6): p. 1490-1507.
114. Kim, Y.J., et al., *pH shock induces overexpression of regulatory and biosynthetic genes for actinorhodin productionin Streptomyces coelicolor A3 (2).* Applied microbiology …, 2007.
115. Lee, E.J., et al., *Identification of sigmaB-dependent promoters using consensus-directed search of Streptomyces coelicolor genome.* J Microbiol, 2004. **42**(2): p. 147-51.
116. Rui, W., et al., *Identification of two-component system AfsQ1/Q2 regulon and its cross-regulation with GlnR in Streptomyces coelicolor.* Molecular Microbiology, 2013. **87**(1): p. 30-48.
117. Rodríguez-García, A., et al., *Arginine boxes and theargRgene inStreptomyces clavuligerus: evidence for a clear regulation of the arginine pathway.* Molecular Microbiology, 1997. **25**(02): p. 219-228.
118. Rosario, P.-R., et al., *ArgR of Streptomyces coelicolor Is a Versatile Regulator.* PLoS ONE, 2012. **7**(3).
119. Smith, C.P. and K.F. Chater, *Structure and regulation of controlling sequences for the Streptomyces coelicolor glycerol operon.* J Mol Biol, 1988. **204**(3): p. 569-80.
120. Chris, D.D.H., et al., *Genes essential for morphological development and antibiotic production in Streptomyces coelicolor are targets of BldD during vegetative growth.* Molecular Microbiology, 2010. **78**(2): p. 361-379.
121. Paula, Y., et al., *Transcriptomic Analysis of Streptomyces coelicolor Differentiation in Solid Sporulating Cultures: First Compartmentalized and Second Multinucleated Mycelia Have Different and Distinctive Transcriptomes.* PLoS ONE, 2013. **8**(3).
122. Yamanaka, K., M. Inouye, and S. Inouye, *Identification and characterization of five cspA homologous genes from Myxococcus xanthus.* Biochim Biophys Acta, 1999. **1447**(2-3): p. 357-65.
123. Sébastien, R., et al., *Feast or famine: the global regulator DasR links nutrient stress to antibiotic production by Streptomyces.* EMBO reports, 2008. **9**(7): p. 670-675.

124. Luo, Y., et al., *Systematic Identification of a Panel of Strong Constitutive Promoters from Streptomyces albus.* ACS Synth Biol, 2015.
125. Dalebroux, Z.D. and M.S. Swanson, *ppGpp: magic beyond RNA polymerase.* Nat Rev Microbiol, 2012. **10**(3): p. 203-12.
126. Arrhenius, S., *Quantitative laws in biological chemistry.* 1915.